EFFECT OF FRAMEWORK AND COMPLEMENTARITY DETERMINING REGION CHARGE ON VH DOMAIN SOLUBILITY

A Thesis Submitted to the College of
Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Master of Science
In the Department of Biochemistry
University of Saskatchewan
Saskatoon

By
Sharmin Sultana

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of the material in this thesis in whole or part should be addressed to:

Head of the Department of Biochemistry
University of Saskatchewan
Saskatoon, Saskatchewan, S7N 5E5

ABSTRACT

The use of VH domains in diagnostic and therapeutic applications circumvents many of the issues associated with antibodies. Nevertheless, VH domains still have limitations such as a tendency to aggregate. They usually contain hydrophobic residues within their complementarity determining regions (CDRs) that facilitate binding to target antigens but also mediate VH domain aggregation, which is a great concern for therapeutic applications. This thesis focuses on the engineering of the VH framework and CDR1 to prevent VH aggregation.

We used two strategies to increase the stability and solubility of the VH domain. First, we used a previously discovered autonomous human VH-B1a as a starting point and increased its net charge by introducing charge mutations in the framework region. We observed that our designed +VH domain showed increased thermostability, but the reversible folding ability and solubility were lower compared to near-neutrally charged VH-B1a. Second, we mutated CDR1 of VH-B1a and +VH domains by introducing a series of positively and negatively charged residues, and assessed the effect of net framework charge and CDR1 charge on improving VH domain thermostability and solubility. From the analysis of CDR1-mutated VH-B1a and +VH domains, we noticed a position specific effect of a single lysine or two lysines in CDR1. Moreover, +VH can overcome the position specific effect of single lysine mutation, and the positional effect of double lysine mutation was better tolerated in +VH domain compared to VH-B1a. Significant aggregation resistance of VH domain was noticed in the presence of three negative (DDD) or positive (KKK) charge in the CDR1 of VH-B1a and +VH, respectively. Furthermore, VH-B1a was better at tolerating negative charge in CDR1, whereas +VH was better at tolerating either positive or negative charge. Our findings demonstrate that in addition to the net framework charge and CDR1 charge, a positional effect of CDR1 mutation affects thermostability and aggregation resistance. Our results demonstrate that the +VH domain will be a good starting template for the construction of phage displayed libraries that will be enriched in highly thermostable and aggregation resistant VH variants.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my thesis supervisor Dr. Ron Geyer for all his invaluable guideline, support, and encouragement. I appreciate his seriousness and guidance throughout my Master's program that makes my research project smooth and interesting. It has been my great pleasure and honor to get a chance to work with him. I would like to thank all of my lab members for their help and support.

I want to acknowledge PCCF at university of Saskatchewan, where I did Circular Dichroism (CD) experiment.

I would like to thank my committee members, Dr. Jeremy Lee, Dr. Miroslaw Cygler, Dr. Scott Napper and Dr. Maruti Uppalapati for their helpful comments and suggestions.

I would also like to express my gratitude to all the faculty, staff and graduate students in the Department of Biochemistry of the University of Saskatchewan for their support during my study.

Finally, I would like to thank my family and friends for their encouragement and support, especially my parents who have devoted all their efforts to me. I want to give special thanks to my husband Nazmul Hasan and daughter Nabiha Samara for their countless support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2YT	2x Yeast extract and Tryptone Broth
Amp	Ampicillin
ATP	Adenosine triphosphate
bp	Base pairs
Carb	Carbenicillin
CCC-dsDNA	Covalently Closed Circular dsDNA
CDRs	Complementarily determining regions
CH	Constant heavy domain
CL	Constant light domain
Cmp	Chloramphenicol
ddH_2O	Sterile double distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	1, 4-dithiothreitol
du-ssDNA	Uracil containing ssDNA
E. coli	*Escherichia coli*
f1 ori	Filamentous phage origin of replication
Fab	Antigen binding fragment
Fc	Fragment crystallizable domain
IgG	Immunoglobin G
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan	Kanamycin
kDa	Kilodalton
LacZ	Beta-galactosidase
LB	Lysogeny broth
NEB	New England Biolabs
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PIII	M13 bacteriophage gene-3 minor coat protein
scFv	Single chain variable fragment
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SOC	Super Optimal Broth with Catabolic repressor medium
Taq	*Thermus aquaticus*
Tm	Melting temperature
Tris	Tris (hydroxymethyl) aminomethane
Tween 20	Polyoxyethylene (20) sorbitan monolaurate
VH	Variable heavy domain
VL	Variable light domain

1. INTRODUCTION

Antibodies are being utilized as effective therapeutics in the clinic. Sales of recombinant therapeutic proteins and antibodies surpassed the US $100 billion in 2010 with therapeutic monoclonal antibodies (mAbs) comprising almost half (48%) of this. At the beginning of July 2011, more than 1000 modified proteins were authorized for medical use in Europe and USA (Dimitrov, 2012). Around 30 monoclonal antibodies were recently approved for therapeutic purposes in Europe, USA, and China, however, a limited number of these antibodies were used as antibody fragments. The principle motivation behind using Immunoglobulin G (IgGs) is the effector function of the fragment crystallizable (Fc) domain, which makes it preferable for therapeutic applications relative to antibody fragments (Dimitrov, 2012).

Immunologists started using antibody fragments for diagnostic and therapeutic purposes in the late 1950s after realizing the bi-functional behavior of IgGs (Porter, 1958, 1959). The fragment antigen-binding (Fab) region of an IgG bind to antigens and the Fc domain mediates the cytotoxic effector functions through antibody-dependent cellular cytotoxicity (ADCC) and/or complement-dependent cytotoxicity (CDC) (Greenwood *et al.*, 1993). Moreover, the Fc domain facilitates the recycling and catabolism of IgGs, which regulates their serum half-life and clearance (Ghetie *et al.*, 1996; Vincke and Muyldermans, 2012). In the early 90s, Inbar and coworkers (Inbar *et al.*, 1972) discovered the variable fragment (Fv fragment) of an IgG through proteolytic digestion, which consists of the variable region of heavy (VH) and light chains (VL). The Fv fragment is produced as a single-chain variable fragment (scFv), where VH and VL domains are linked through a Gly-rich short peptide linker (Huston *et al.*, 1991). The generation of scFv fragments provided the idea to design VH and VL nanobodies, which are single-domain antibody fragments. The antigen-binding properties of camelids' VHH domains, which are devoid of light chains, brought about the idea to use nanobodies as therapeutics. Nanobodies are a useful diagnostic and therapeutic tools because of their remarkable biophysical properties of high stability, solubility, and strict monomeric behavior, which makes them particularly attractive for clinical use (Conrath *et al.*, 2001). Additionally, due to the small size of nanobodies, they can easily penetrate the tumor and be rapidly cleared *in vivo*, which makes the VH domain suitable to deliver toxins for therapy and radioisotopes for imaging (Rosebrough, 1993; Zuckier and DeNardo, 1997; Cortez-Retamozo *et al.*, 2002; Sundaresan *et al.*, 2003; Barthelemy *et al.*, 2008). Unfortunately, isolated VH or VL domains

have a large hydrophobic surface that is exposed to the external aqueous environment, which causes them to self-associate and aggregate under certain conditions. These limitations can be overcome by engineering the VH/VL interface, or the VH or VL domain itself (Vincke and Muyldermans, 2012).

2. LITERATURE REVIEW

2.1. Antibodies

The immune system provides a central defense against pathogenic organisms and toxins by producing proteins called antibodies. Antibodies show high binding affinity and specificity towards their target molecules (antigens) and are widely used in research, diagnosis, and therapy (Perchiacca *et al.*, 2012). They are attractive therapeutic molecules because of their specificity, long *in vivo* half-life (typically two to four weeks), and lack of immunogenicity (Tan *et al.*, 2002; Hwang and Foote, 2005; Zalevsky *et al.*, 2010).

In humans, there are five antibody isotypes referred to as IgA, IgD, IgE, IgG, and IgM. IgG is the most common class of antibody, which is a heterotetramer composed of two identical pairs of light (25 kDa) and heavy (50 kDa) chains linked by interchain disulfide bonds (Padlan, 1994) (Figure 2.1A). The heavy chain is comprised of one variable heavy (VH), and three constant heavy domains (CH1, CH2, and CH3), whereas the light chain consists of one constant (CL) and one variable domain (VL). The antigen-binding fragment (Fab) is comprised of the CH1, CL, VH, and VL domains and the antigen binding site is located on VH and VL domains. Variable heavy (VH) and variable light (VL) domains each contain three hypervariable loops, called complementarity determining region (CDRs), which are responsible for antigen binding. Furthermore, effector function is facilitated by the Fc domain of antibody comprised of CH2 and CH3 domains (Tonikian and Sidhu, 2012).

Since the binding activity of an antibody is localized to its variable domain it can be genetically engineered to retain the binding activity of its parental antibody. Examples of genetically engineered antibody fragments include antigen-binding fragment (Fab), single chain variable fragment (scFv), variable heavy chain domain (VH), and variable light chain domain (VL). Fabs are the most widely used engineered antibody fragments because they contain both variable domains and stabilizing constant domains; whereas in scFvs, VH and VL domains are connected via a flexible peptide linker to limit the dissociation of the two variable domains. Single variable domains, such as VH or VL nanobodies can be engineered so that they can bind to diverse targets with high affinity (Figure 2.1B) (Bird *et al.*, 1988; Huston *et al.*, 1988; Ward *et al.*, 1989; Holliger *et al.*, 1993; Davies and Riechmann, 1995; Plückthun and Pack, 1997;

Decanniere *et al.*, 1999; Muyldermans, 2001; Holt *et al.*, 2003; Feldhaus and Siegel, 2004; Sidhu *et al.*, 2004).

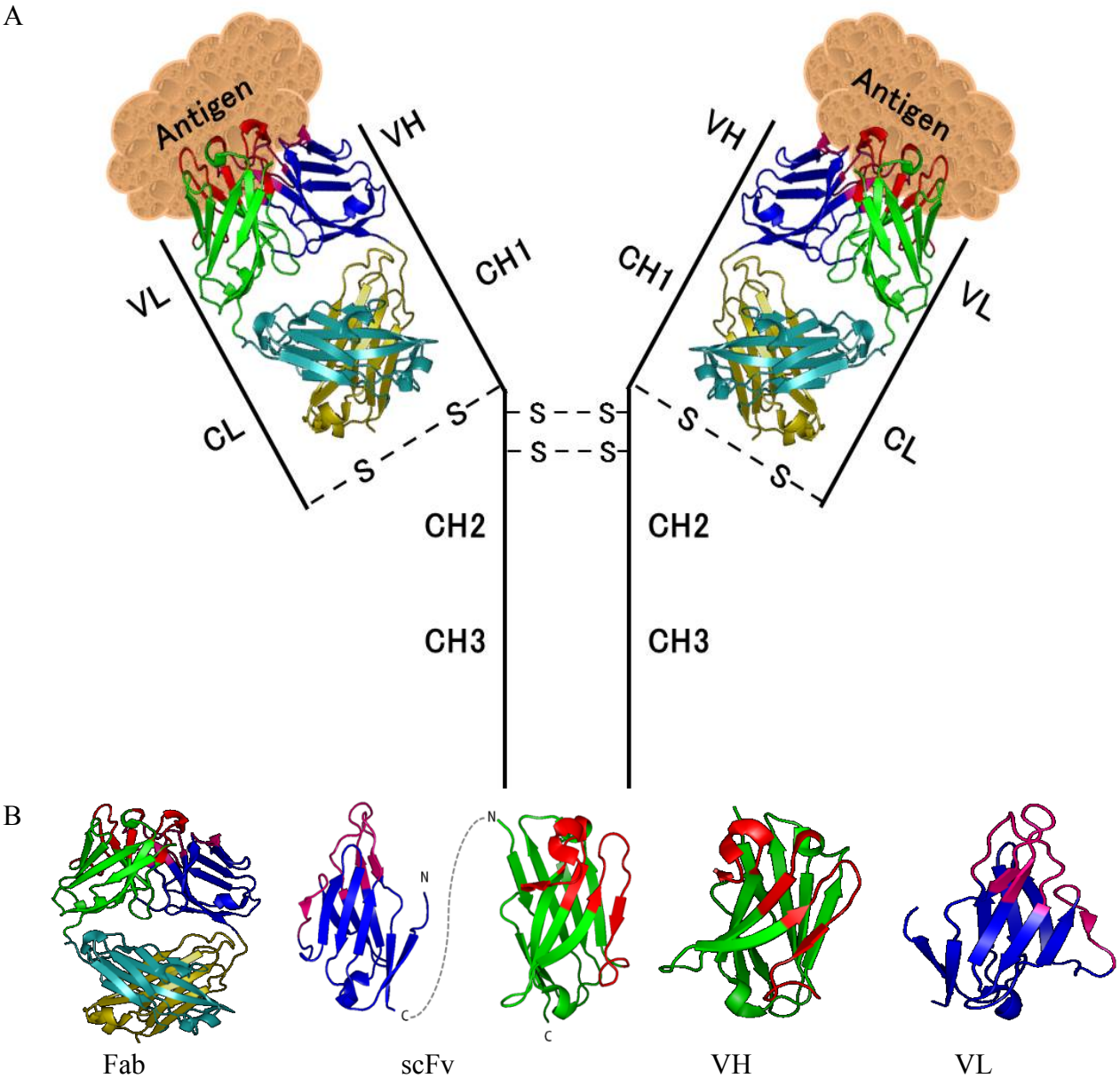


Figure 2.1. Molecular architecture of IgG and genetically engineered antibody fragments
(A) Schematic of the heterotetrameric IgG molecule composed of two light chains and two heavy chains. The heavy chain is comprised of one variable heavy (VH) and three constant heavy (CH) domains. The smaller light chain consists of one variable light (VL) and one constant light (CL) domain. Interchain disulfide bonds are shown as (-S-S-). (B) The antigen-binding fragment (Fab) is comprised of the CH1, CL, VH, and VL domains and is depicted as a cartoon (PDB ID: 3DVG). Variable heavy (VH) and variable light (VL) domains each contain three hypervariable loops known as CDRs that bind to the antigen and shown in red and pink, respectively. Fabs can be further reduced to scFvs, where VH and VL domains are linked via a peptide linker. VH and VL are the single antigen-binding domain of the antibody.

2.2. Single domain antibodies

Monoclonal antibodies (mAbs) are well established as therapeutic drugs because of their high potency, high specificity, tolerance, and long half-life in serum. However, they possess some properties that do not make them desirable for all applications. For example, their large size (~150 kDa) confines them to the extracellular environment and reduces their ability to penetrate solid tumors. They can also recruit unwanted effector functions through their Fc domain (Ignatovich *et al.*, 2012). Moreover, expression of antibodies in the mammalian cell is expensive and time-consuming, and expression using less expensive systems such as bacteria is challenging for a number of reasons. First, folding and assembly of the heterotetramer IgGs occur after their translocation through the endoplasmic reticulum membrane, which is assisted by several molecular chaperones such as BiP and PDI, and this mechanism is not possible using bacterial expression system (Baral and Arbabi-Ghahroudi, 2012). Second, heavy and light chains are linked together by several disulfide bonds, which do not form in the reducing environment of bacterial cytoplasm. Third, post-translational modification, i.e. glycosylation of the heavy chain, is required for biological activities of antibodies, which is not possible in bacteria since they lack the appropriate glycosylation machinery. Further, antibodies lacking glycosylation can be improperly folded or degraded by the cell (Andersen and Reilly, 2004; Baral and Arbabi-Ghahroudi, 2012). Therefore, mammalian cell (Chinese hamster ovary (CHO) and NS0 cells) expression systems are most often used to produce glycosylated antibodies. Unfortunately, mammalian cell expression of recombinant antibody is costly and labor-intensive (Verma *et al.*, 1998). It takes approximately 4-6 months to manufacture a stable mammalian cell line production system, whereas for bacteria it takes one month. Third, it becomes difficult to produce antibodies in sufficient quantity and quality for basic research and preclinical studies (Andersen and Reilly, 2004).

Due to these limitations in producing and using antibodies, researchers are interested in single domain antibodies (sdAbs). It is well established that many problems related to *in vivo* expression, correct folding, solubility, thermostability, and conformational stability of antibodies can be avoided by reducing their complexity and the size (Baral and Arbabi-Ghahroudi, 2012). In this regard, different antibody fragment formats like scFvs, VH, and VL domains have been produced that are well expressed in bacterial systems. Moreover, the small size of VH domains (~13 kDa) facilitates their increased clearance *in vivo* and enhanced tumor

penetration, which is challenging for an antibody molecule. Therefore, VH domains are now being used for delivering drugs or toxins to any target and can be employed for nuclear and optical imaging (Rosebrough, 1993; Zuckier and DeNardo, 1997; Cortez-Retamozo *et al.*, 2002; Sundaresan *et al.*, 2003; Barthelemy *et al.*, 2008).

2.3. Heavy chain variable domain (VH) antibody fragment

The VH domain is small in size (~13 kDa) and consists of a sandwich of two antiparallel β-sheets connected by a conserved intramolecular disulfide bond (Padlan, 1994; Vincke and Muyldermans, 2012). The VH domain contains three complementarity determining regions: CDR1, CDR2, and CDR3, which are highly variable in length and sequence. They are involved in antigen recognition through non-covalent interactions (Figure 2.2). CDRs are separated by framework regions that act as a scaffold to support and display them.

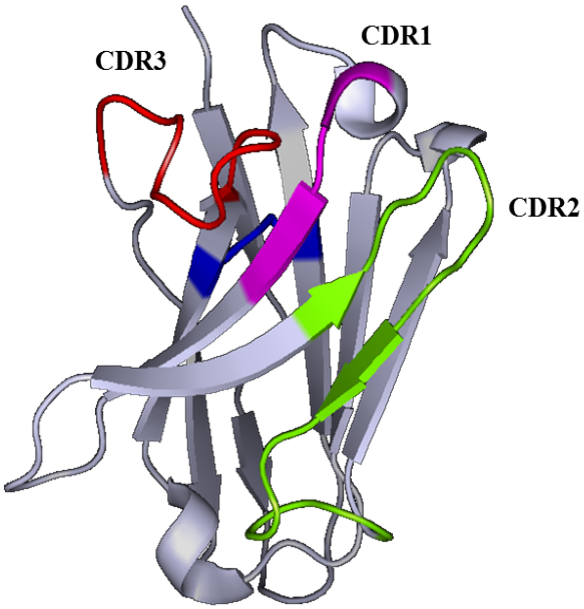


Figure 2.2. VH domain structure
Crystal structure of human VH domain (PDB: 3B9V). CDR1 is highlighted in pink, CDR2 in green, and CDR3 in red. Blue bond denotes a disulfide bond between cysteine residues at position 22 and 92.

The main limitation of using the VH domain is its propensity to self-associate and aggregate under various stressed conditions such as high concentrations, elevated temperatures, freeze-thaw cycles, agitation, low pH, and long storage times (>2 years for therapeutic

applications) (Shire *et al.*, 2004; Shire, 2009). In antibodies, VH and VL domains interact with each other through an interface composed of hydrophobic residues (Perchiacca and Tessier, 2012). These hydrophobic residues are exposed to the external environment when the VH domain is expressed alone, leading to self-association and in some cases aggregation through hydrophobic interactions (Barthelemy *et al.*, 2008; Perchiacca *et al.*, 2012; Tonikian and Sidhu, 2012). VH domain aggregation is of particular concern for therapeutic applications because such aggregates can be immunogenic (Perchiacca and Tessier, 2012).

2.3.1. VH domain optimization

Since the entire antibody is not required for antigen binding, antibody fragments have been engineered to circumvent many of the issues associated with antibodies (Figure 2.3A). Interest in developing human VH domains was stimulated by the discovery of autonomous VH domains in camelids (camels, dromedaries, and llamas) and VH-like domains in cartilaginous fish (wobbegong and nurse sharks) (Saerens *et al.*, 2008b).

Interestingly, in humans the presence of "heavy chain antibodies" (HCAbs) are related to a pathological disorder, known as the heavy chain disease and occurs in sera of patients. These truncated antibodies are generated from a somatic event that removes various parts of the VH and CH1 region from the expressed IgG gene. Truncated human HCAbs are not functional for antigen binding since the VL and part of the VH domain are missing (Alexander *et al.*, 1982; Vincke and Muyldermans, 2012).

Camelids produce "heavy chain antibodies" (HCAbs) that are devoid of light chains, termed VHH (heavy chain variable domain of a heavy chain antibody) (Hamers-Casterman *et al.*, 1993) (Figure 2.3B). In contrast with human HCAbs isolated from patients with heavy chain disease, camelid HCAbs contribute to the immune response of these animals (Spinelli *et al.*, 1996). Recombinant expression of the VHH yields a soluble single domain antibody (sdAb) of 15 kDa that is autonomously stable in the absence of a light chain partner. Furthermore, as VHH domains are relatively small, they exhibit increased clearance *in vivo* and enhanced tumor penetration. Thus, they are preferable for specialized therapeutic applications (Cortez-Retamozo *et al.*, 2002). They are distinct from conventional VH domains by the substitution of five amino acids (L12S, V37F/Y, G44E, L45R/C, and W47G) that are conserved in all VH domains of vertebrate antibodies (Davies and Riechmann, 1994). The substitution of

hydrophobic Leucine with hydrophilic Serine at position 12 increases the solubility of VHHs in an aqueous environment. Moreover, residues at positions 42, 49, 50, and 52 are part of the VH/VL interface, and substitution of non-polar with polar amino acids (G49E and L50R) increases the hydrophilicity of the VHH surface. Furthermore, substitutions at positions 42 and 52 cause a net shift of the bulky hydrophobic groups towards the center of the five-stranded β-sheet and expose their most hydrophilic parts to the solvent. Since the VH/VL interface of camelid VHH is hydrophilic, the VHH domain cannot associate with a VL domain (Howard and Kaser, 2007). Consequently, Camelid VHH domain is stable and soluble in the absence of VL domain.

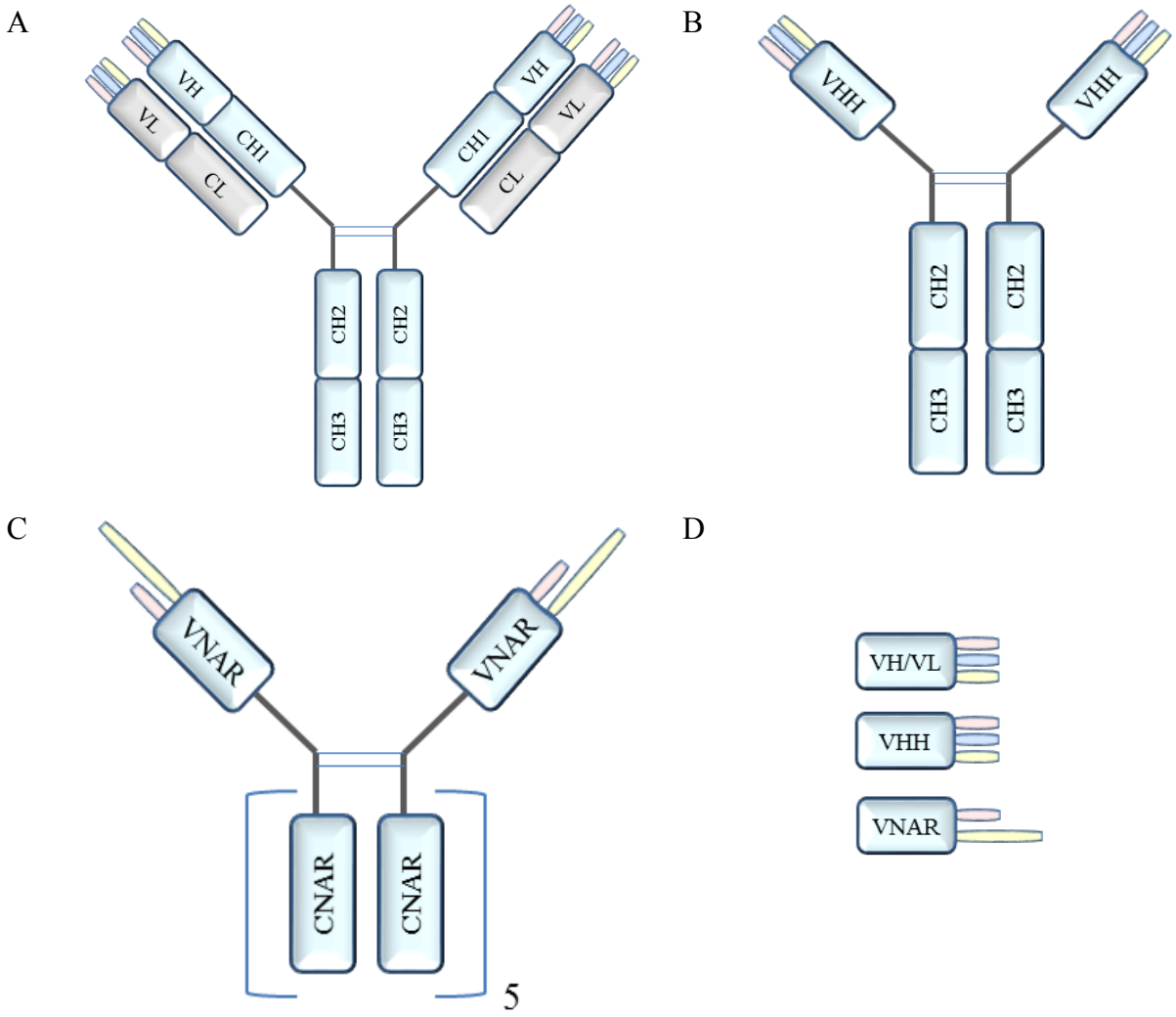


Figure 2.3. Structure of conventional IgG, camelid HCAb, shark HCAb, and nanobodies
Conventional IgG (A), camelid HCAb (B), shark HCAb (C), and nanobodies made of VH or VL domains, camelid VH domains (VHH), and shark VH domains (VNAR) (D). The red bar represents CDR1, blue bar CDR2, and yellow bar represents CDR3.

The third type of heavy chain antibody called "Immunoglobulin new antigen receptors" (IgNARs) are found naturally in sharks (Streltsov *et al.*, 2005; Nuttall, 2012). They are homodimers containing one variable and five constant domains with no associated light chains. Antigen binding ability of IgNAR is located on the variable domains called VNARs, which are ~13 kDa. VNARs exhibit significant variability within CDR1 and CDR3 loops, and residues critical for antibody VH-VL dimerization are mutated or missing (Greenberg *et al.*, 1995). Conventional antibodies bind antigen using up to six CDRs (Chothia *et al.*, 1989; Padlan, 1994), whereas VNARs recognizes antigen using two CDRs (Roux *et al.*, 1998; Nuttall *et al.*, 2003). To compensate for their reduced number of CDRs, IgNARs encode unusually long, structurally complex and highly variable CDR3s (Figure 2.3C) (Greenberg *et al.*, 1995; Nuttall *et al.*, 2004). By observing all these naturally occurring VH domains, nanobodies can be made using VH or VL domains from an IgG, camelid VHH domains, and shark VNAR domains (Figure 2.3D).

2.3.2. Applications of VH domains

Monoclonal antibodies are not optimal for many applications because of their large size, which limits their effective penetration into solid tumors and the passage through the blood-brain barrier. This limitation of monoclonal antibodies inspired the development of recombinant antibodies (scFv, VHH, VH, and VL) of smaller size. Recently, VH domains have been broadly used in biotechnological, therapeutic, and diagnostic applications due to their small size, simple structure, high stability, natural production, and high affinity for the target (de Marco, 2011; Muyldermans, 2013).

2.3.2.1. VH domain in biotechnological applications

The VH domain offers several advantages for biotechnological applications due to their single domain structure (Vincke and Muyldermans, 2012). They can be used to detect disease-associated proteins, study protein-protein interactions, and facilitate protein crystallization. Camelid heavy chain antibodies (VHH) can be fused to fluorescent proteins to generate chromobodies that can recognize and track targets in different compartments of living cells. Chromobodies allow live cell microscopy and enable novel functional studies (Rothbauer *et al.*, 2006). For example, chromobodies are used to visualize the gankyrin oncoprotein in cells by

labeling VHH domains with enhanced green fluorescent protein (eGFP) and red fluorescent protein (RFP). Once VHH domains interact with gankyrin on different epitopes, the gankyrin can be detected using fluorescence lifetime imaging (FLIM)-fluorescence resonance energy transfer (FRET) microscopy (Rinaldi *et al.*, 2013; Liu and Geyer, 2015). An additional application is through GFP-nanotraps, which is a GFP-binding protein that is formed from the combination of an anti-GFP nanobody to a monovalent matrix, such as agarose beads, magnetic agarose beads, or magnetic particles. It can be used for the biochemical analysis of GFP fusion proteins and their interacting partners (Rothbauer *et al.*, 2008). Nuclear imaging also employs nanobodies using a technique to examine the structure and function of any target organ or protein, where a specific delivery of fluorophores into proximity with the target is mandatory. Due to the large size of full-length antibodies, the distance between the fluorophore and the target is higher, producing a misleading signal. However, fluorophore-labeled anti-GFP nanobodies allow nanometer spatial resolution and minimal linkage error because of the high affinity and small size of nanobodies (Ries *et al.*, 2012). Additionally, camelid VHHs can be used as crystallization chaperones to increase the crystallization probability of a target molecule. They reduce the conformational heterogeneity of a target molecule by binding and stabilizing its conformational states. Moreover, they mask the surface that interferes with the crystallization process and extends regions that form crystal contacts (Koide, 2009).

2.3.2.2. VH domain in diagnostic applications

VH domains can be used as affinity agents in biosensor applications. Due to their small size, VH domains can be linked more densely on biosensor surfaces and they can target less accessible epitopes, such as in tumors. Furthermore, their conformational stability leads to higher resistance to surface regeneration conditions (Saerens *et al.*, 2005; Deffar *et al.*, 2009). Recently camelid VHH domains have been used in cancer diagnosis, such as early detection and staging of prostate cancer by detecting different isoforms of prostate-specific antigen (PSA) in the blood (Saerens *et al.*, 2004). In 2008, Lieven Huang and his team used llama single domain antibody fragment for the radio-immunodetection of EGFR over-expressing tumors in mice. They used 99mTc-8B6 nanobody to discriminate between high and moderate EGFR over-expressing tumors using single photon emission computed tomography (SPECT) (Huang *et al.*, 2008). Nanobodies can also be exploited in the detection of the larval stage of pork

tapeworm (*Taenia solium*), which causes Neurocysticercosis (NCC). The condition can be diagnosed using nanobodies generated against the *Taenia solium* (Deckers *et al.*, 2009).

2.3.2.3. VH domain in therapeutic applications

VH domains can be used for the treatment of various diseases because of their ability to target toxic enzymes and block specific molecular interactions. For example, Trypanosomiasis, or sleeping sickness was treated successfully with VHH domains fused to the apolipoprotein L-1 (ApoLI) enzyme. This enzyme can bind to the trypanosome coat protein and lyse the parasite (Stijlemans *et al.*, 2004; Baral *et al.*, 2006; Deffar *et al.*, 2009). In oncology, lactamase conjugated carcinoembryonic antigen (CEA)-specific VHH was used for targeting tumor cells. This lactamase enzyme kills the tumor cells by converting an injected non-toxic prodrug into a toxic drug in the neighborhood of the targeted tumor cells (Cortez-Retamozo *et al.*, 2004). Camelid VHHs have been generated against epidermal growth factor receptor (EGFR) that block the binding of epidermal growth factor (EGF) with its receptor and prevent tumor formation (Roovers *et al.*, 2007). Furthermore, camelid VHHs that bind to the bovine tumor necrosis factor can be utilized for the treatment of Crohn's disease and rheumatoid arthritis (Maass *et al.*, 2007). In 2008, a new VHH nanobody was discovered that could recognize and neutralize Androctonus australis hector AahI' toxin, which is a serious public health problem in many countries (Hmila *et al.*, 2008). Various gastrointestinal tract diseases, such as inflammatory bowel disease and colon cancer can be treated with orally available single domain antibodies. Oral application of the llama single-domain antibody fragment (VHH) reduces *E. coli*-induced diarrhea in piglets (Harmsen *et al.*, 2006). Moreover, camelid VHH can be used in photothermal therapy for the treatment of HER2-induced breast and ovarian cancer. Anti-HER2 VHHs were conjugated with branched gold nanoparticles. Electromagnetic radiation was used to heat up the gold nanoparticles that burn the cancer cell membrane at the bound spot thereby killing cancer cells (Van de Broek *et al.*, 2011). ALX-0081 is a single-domain antibody that targets von Willebrand factor and prevents thrombosis formation in patients with acute coronary syndrome (Bartunek *et al.*, 2013).

2.4. Strategies to design aggregation-resistant human VH domains

A drawback of the human VH domain is its poor stability, which is one of the key determinants of antibody efficacy. Human VH domains have been engineered to be as stable as larger antibody fragments and monoclonal antibodies. A number of strategies have been employed to improve VH domain thermostability and solubility without reducing target binding affinity. These include disulfide bond insertion, camelization approach**,** charge engineering, and random mutagenesis (Perchiacca *et al.*, 2012; Perchiacca and Tessier, 2012).

2.4.1. Engineering disulfide-bonds to enhance VH domain stability

VH domain stability can be increased by introducing a second disulfide bond between the two opposing β-sheets (Hagihara *et al.*, 2007; Chan *et al.*, 2008; Saerens *et al.*, 2008a; Hussack *et al.*, 2011; Hussack *et al.*, 2012; Kim *et al.*, 2012; Lee *et al.*, 2013). Recently, several camelid VHHs against human prostate-specific antigen (hPSA) have been isolated and characterized (Saerens *et al.*, 2004). Two similar VHH domains specific for hPSA (i.e., cAbPSA-N8 and cAbPSA-C12) have identical antigen binding loop sequences and similar kinetic binding parameters. The cAbPSA-C12 VHH possesses two Cys residues at positions 54 and 78 of its framework regions; whereas a homologous clone cAbPSA-N8 carries Gly and Ile instead of Cys. The disulfide bond of cAbPSA-C12 increased its thermal and conformational stabilities compared to the cAbPSA-N8 without Cys residues at those positions. This observation gave rise to a strategy of stabilizing single domain antibodies by introducing disulfide bond between two β-sheets (Saerens *et al.*, 2008a). In camels, VHHs typically have only one intradomain disulfide bond. An additional intradomain disulfide bond can be introduced by substituting two amino acids with Cys residues at position 49 and 69 (numbering according to Kabat *et al.* (Kabat *et al.*, 1992)). These substitutions have minimal effects on antigen binding affinity, but significantly increase the thermodynamic stability compare to the wild type VHH (Hagihara *et al.*, 2007; Saerens *et al.*, 2008a; Hussack *et al.*, 2011; Hagihara and Saerens, 2014; Walper *et al.*, 2014; Akazawa-Ogawa *et al.*, 2016). Kim *et al*. used four models of human VH domains and established that like camelid VHH domains, the introduction of a second disulfide bond increases the thermostability of human VH domains (Kim *et al.*, 2012). The disulfide linkage between Cys23 and Cys104 residues in the core of VH domains is highly conserved and crucial for their stability and function (Goto and Hamaguchi,

1979; Proba *et al.*, 1998; Ciaccio and Laurence, 2009). Moreover, the addition of an extra disulfide linkage between Cys54 and Cys78 residues increases the thermostability of human VHs by 14-18°C and also reduces VH aggregation (Wetzel *et al.*, 1988; Matsumura *et al.*, 1989; Young *et al.*, 1995; Davies and Riechmann, 1996; Mansfeld *et al.*, 1997; Hagihara *et al.*, 2007; Chan *et al.*, 2008; Saerens *et al.*, 2008a; Gong *et al.*, 2009; Hussack *et al.*, 2011; Govaert *et al.*, 2012; Wozniak-Knopp *et al.*, 2012). The disulfide linkage stabilizes VH domains by reducing their conformational entropy in unfolded states (Fersht, 1997; Mason *et al.*, 2002). However, due to the conformational change caused by the engineered disulfide linkage these mutant VH domains show reduced binding to Protein A, which is used for the purification of correctly folded VH domain (Kim *et al.*, 2012; Lee *et al.*, 2013). Engineering disulfide bonds into VH domains does not always enhance its stability due to unfavorable effects of the surrounding amino acids of the folded VH domain (Betz, 1993) or replacement of an existing favorable interaction. Another reason for the limited success of disulfide bond engineering is due to the lack of firmly established rules for selecting appropriate loci for disulfide cross-linking (Saerens *et al.*, 2008a).

2.4.2. Camelization approach to enhance VH domain stability

Camelidae family members produce stable VHH domains that are devoid of light chains, which has given insight into the design of stable human VH domains (Hamers-Casterman *et al.*, 1993; Barthelemy *et al.*, 2008). Camelid VHH domains are thermostable and can fold reversibly after thermo-denaturation (Ewert *et al.*, 2002). The VH/VL interface of camelid VHH domain contains charged amino acids instead of hydrophobic residues of the VH/VL interface of human VH domain (Desmyter *et al.*, 1996; Spinelli *et al.*, 1996; Decanniere *et al.*, 1999; Spinelli *et al.*, 2000; Desmyter *et al.*, 2001; Spinelli *et al.*, 2001; Desmyter *et al.*, 2002). Moreover, the largest differences between VHH and VH domains occurs at four positions of the second framework region (Phe37, Glu44, Arg45 and Gly47), referred to as the "VHH tetrad" (Harmsen *et al.*, 2000; Nguyen *et al.*, 2000; Muyldermans *et al.*, 2001; Holt *et al.*, 2003). This VHH tetrad decreases the hydrophobicity of VH/VL interface of the camelid VHH domain. Consequently, camelid VHH domain can unfold and refold reversibly into a functional form after denaturation that enhances their solubility and stability (Desmyter *et al.*, 1996; Spinelli *et al.*, 1996; Decanniere *et al.*, 1999; Spinelli *et al.*, 2000;

Desmyter *et al.*, 2001; Perez *et al.*, 2001; Spinelli *et al.*, 2001; Desmyter *et al.*, 2002; Ewert *et al.*, 2002). In addition, Camelid VHHs have a longer CDR3 (average of 18 amino acids) compared to human VH domains (average of 14 amino acids) (Muyldermans *et al.*, 1994), which pack against hydrophobic regions of the VHH domain analogous to the VH/VL interface of human VH domains (Desmyter *et al.*, 1996; Harmsen *et al.*, 2000; Desmyter *et al.*, 2002), thus sequestering hydrophobic residues from VH/VL interface. The presence of charged residues at the VH/VL interface and the packing of extended CDR3 loops against hydrophobic interface contributes to the stability and high level of solubility of the camelid VHH domain. Attempts have been made to construct soluble human VH domains by introducing elements of camelid VHH tetrads into the human VH domains, a process referred to as "camelization approach" (Davies and Riechmann, 1994; Desmyter *et al.*, 1996; Spinelli *et al.*, 1996; Decanniere *et al.*, 1999; Spinelli *et al.*, 2000; Desmyter *et al.*, 2001; Spinelli *et al.*, 2001; Desmyter *et al.*, 2002). The introduction of these changes into human VH domains has brought modest improvements in the solubility and thermostability of human VH domain. On the other hand, it appears that camelization strategies have significant constraints on the design of CDR3 randomized phage-displayed VH domain libraries. Since the VH domain mainly interacts with antigen through its CDR3 loop, the CDR3 diversity is critical for the construction of phage-displayed VH libraries (Sidhu *et al.*, 2000; Lee *et al.*, 2004; Howard and Kaser, 2007).

2.4.3. Influence of CDR loops on VH domain stability

Antibodies commonly contain hydrophobic residues within their CDRs that mediate binding to target antigens. Unfortunately, these hydrophobic CDRs interact with hydrophobic residues of unfolded antibodies and promote antibody aggregation (Mian *et al.*, 1991; Padlan, 1994; Davies and Cohen, 1996; Sundberg *et al.*, 2000; Fellouse *et al.*, 2005, 2006; Perchiacca *et al.*, 2012). The enhanced folding stability of a human VH domain due to mutations at edges of their CDR loops suggests that the sequence of CDRs may have an impact on antibody folding stability (Helms and Wetzel, 1995; Ionescu *et al.*, 2008). Moreover, the ability of human VH and camelid VHH domains to resist aggregation after unfolding could be determined by the sequence of their CDR loops (Bond *et al.*, 2003; Jespers *et al.*, 2004; Arbabi-Ghahroudi *et al.*, 2009; Vincke *et al.*, 2009; Perchiacca *et al.*, 2011). Jespers and coworkers used phage display to study the role of CDR sequences in VH domain aggregation. They constructed phage

displayed VH domain libraries with random variations in CDR sequences, and VH domains were displayed in a multivalent format on the PIII coat protein of filamentous phage. Phage libraries were heated transiently at 80°C to unfold VH domains and then cooled. They selected aggregation-resistant VH domains using a Protein A-based binding strategy that only binds with correctly folded VH domains. Interestingly, they observed that the only difference between CDR sequences of highly and poorly soluble VH domains is a moderate increase in the number of negatively charged amino acid residues in the highly soluble variants (Jespers *et al.*, 2004). Recently, Perchiacca and his group investigated whether the aggregation resistance of VH domain is determined by a particular CDR sequence or properties of all three CDRs. They grafted corresponding CDRs from aggregation-resistant Hel4 VH domain to the aggregation-prone VH domain and revealed that a triad of three negatively charged amino acids (^{31}DED^{33}) in CDR1 was sufficient to prevent VH domain aggregation. Moreover, phenylalanine to aspartate mutation at position 29 (F29D) that is adjacent to CDR1 also endows reversible folding behavior to the VH domain. Their study shows that mutations within or near the CDR1 remove one aggregation "hot spot" within CDR1 that is responsible for providing aggregation resistance behavior of VH domains (Perchiacca *et al.*, 2011). To examine whether these aggregation resistant mutations could be combined to construct synthetic VH domain libraries, Mandrup and his group used a modified Hel4 sequence as the template for phage displayed VH library construction (Mandrup *et al.*, 2013). They combined aggregation resistant CDR1 of Hel4 with an additional aggregation preventing I29D mutation (Perchiacca *et al.*, 2011) and randomized CDR2 and CDR3 for library construction. They used this library to isolate VH domains against lysozyme and human brain vascular pericytes (HBVP) cells. The isolated VH domains have lower thermodynamic stability than I29D VH domain mutant (Perchiacca *et al.*, 2011) and four out of six isolated VH domains fold reversibly. Their analysis suggests that I29D mutation is not compatible with aggregation resistant CDR1 (^{31}DED^{33}) and indicates that mutation combination that confers improved VH domain stability still needs to be tested empirically (Liu and Geyer, 2015).

Perchiacca and coworkers observed that insertion of two or more negatively charged residues at the edges of CDR3 prevents aggregation of the anti-Alzheimer's amyloid β-peptide VH domain without altering their binding affinity. Further, they found that negatively charged residues need to be positioned closest to the hydrophobic portion of the CDR3 (Perchiacca *et*

al., 2012). In another study, Perchiacca and coworkers showed that the net charge of VH domain determines the optimal CDR3 mutations for preventing VH aggregation. They found that negatively charged CDR3 residues are more efficient for reducing aggregation of VH domains that are negatively or near-neutrally charged, whereas positively charged residues are more suitable for decreasing aggregation of positively charged VH domains (Perchiacca *et al.*, 2014). Moreover, Dudgeon and coworkers observed that aggregation tendency of the VH domain is strongly influenced by the sequences of the CDR1 and negatively charged residues in the CDR1 increase heat-induced aggregation resistance of human VH domain than the positively charged residues. Furthermore, they observed a preference for aspartate over glutamate at all analyzed positions (Dudgeon *et al.*, 2012).

2.4.4. Framework mutations to enhance VH domain stability

To comprehensively analyze factors that are required for the stability and solubility of autonomous human VH domain, Barthelemy and co-workers constructed a phage-displayed VH domain library by randomizing various residues involved in the stabilization of camelid VHH domains as well as in the VH/VL interface (Barthelemy *et al.*, 2008). Protein A-based selection was used to select autonomous VH domains from these libraries and selected VH domains were analyzed using quantitative saturation scanning (Pal *et al.*, 2006) and shotgun alanine scanning methods (Weiss *et al.*, 2000). Interestingly, they identified five mutations (H35G, Q39R, L45E, R50S, and S93A) in three β-strands, which significantly increased the thermostability and reversible folding ability of the VH domain (Barthelemy *et al.*, 2008). The first pair of mutations, Q39R and L45E, are oppositely charged, located near the base of the VH domain, and stabilized the folded structure via a favorable electrostatic interaction. The second pair of mutations, H35G and R50S, are localized at the edges of CDR1 and CDR2, and increase the folding stability of VH domain by changing the orientation of adjacent hydrophobic side chains of aromatic residues (Barthelemy *et al.*, 2008; Perchiacca and Tessier, 2012). These four mutations together provide enhanced thermostability and allow VH domain to fold reversibly.

A VH domain having these mutations plus three additional mutations (W47L, S50R, W103S) was used to construct phage displayed VH domain libraries. These libraries were successfully screened against vascular endothelial growth factor (VEGF) to isolate VH domains with low nanomolar affinity (Ma *et al.*, 2013).

2.4.5. Influence of charge on VH domain stability

Stability and solubility of antibody fragments can be increased by altering their net charge (Jespers *et al.*, 2004; Arbabi-Ghahroudi *et al.*, 2009; Dudgeon *et al.*, 2009; Kvam *et al.*, 2010; Perchiacca *et al.*, 2011, 2012). Miklos and coworkers used the "Rosetta" computational package to evaluate the impact of charge mutations on folding stability of an anti-MS2 scFv. They generated five positively and four negatively "supercharged" scFvs. The antigen-binding activity of negatively charged scFvs was significantly reduced, whereas positively charged scFvs retained significant binding activity after heating at 70°C for 1 hour. Moreover, positively charged scFvs were much more resistant to aggregation than the wild type or negatively charged variants (Miklos *et al.*, 2012). David Liu's research group observed a relationship between net charge and aggregation resistance of protein. They designed positively and negatively supercharged green fluorescent protein (GFP), streptavidin and Glutathione-S-transferase (GST). They observed that both positively (arginine, lysine) and negatively (glutamate, aspartate) charged residues on the solvent-exposed surface of the GFP and streptavidin provide aggregation resistance. They also observed aggregation resistance for highly negatively charged GST, but not for highly positively charged GST (Lawrence *et al.*, 2007). Perchiacca and coworkers observed that the net charge of the VH scaffold determines whether positive or negatively charged residues introduced into CDR3 will prevent VH domain aggregation. They found that negatively charged residues are more efficient in decreasing aggregation of VH domains that are negative or near-neutrally charged, whereas positively charged residues prevent aggregation of positively charged VH domains (Perchiacca *et al.*, 2014).

2.4.6. Rational and random VH domain mutations by phage display

In 1985, George Smith first described the phage display technology for selecting peptides from combinatorial peptide libraries (Smith, 1985). Five years later, it was revealed that this technique could also be used to display antibody fragments (McCafferty *et al.*, 1990). Phage display is a powerful technique that provides a physical linkage between the genotype and phenotype. This linkage is possible because of the use of a specialized vector called a phagemid, which contains the following components: First, the phagemid comprises a dsDNA origin of replication as well as ssDNA filamentous phage origin of replication (f1 ori). The f1

ori allows packaging of the DNA into the phage particles. Second, the phagemid contains a fusion protein consisting of the protein of interest fused to the gene-3 minor coat protein (P3 or pIII) of the filamentous bacteriophage M13 (Lee *et al.*, 2007), which is displayed on the outside surface of the bacteriophage. In this strategy, phage displayed VH domain libraries are screened for binding to target antigen under conditions that allow selection of VH variants with enhanced stability (Jung *et al.*, 1999). Using the phage display method, VH domains that are able to bind to the target are retained and VH domains that do not bind to the target are washed away. Bound phages are eluted and used immediately for infection and amplification in a bacterial host strain. Once purified, the newly enriched phage library can be used for analysis or another round of selection (Figure 2.4).

Jespers and coworkers discovered aggregation-resistant VH domain through phage display method that has negatively charged residues in the CDR loop (Jespers *et al.*, 2004). Using the phage display method, Ghahroudiand and his coworkers observed that an aggregation-resistant VH domain has both negatively charged CDR loop and an extra disulfide bond between CDRs (Arbabi-Ghahroudi *et al.*, 2009). Barthelemy and coworkers discovered an autonomous human VH domain (VH-B1a) through phage display method. Interestingly, they identified five mutations (H35G, Q39R, L45E, R50S, and S93A) in three β-strands, which significantly increase the folding stability and thermostability of VH-B1a (Barthelemy *et al.*, 2008).

A

B

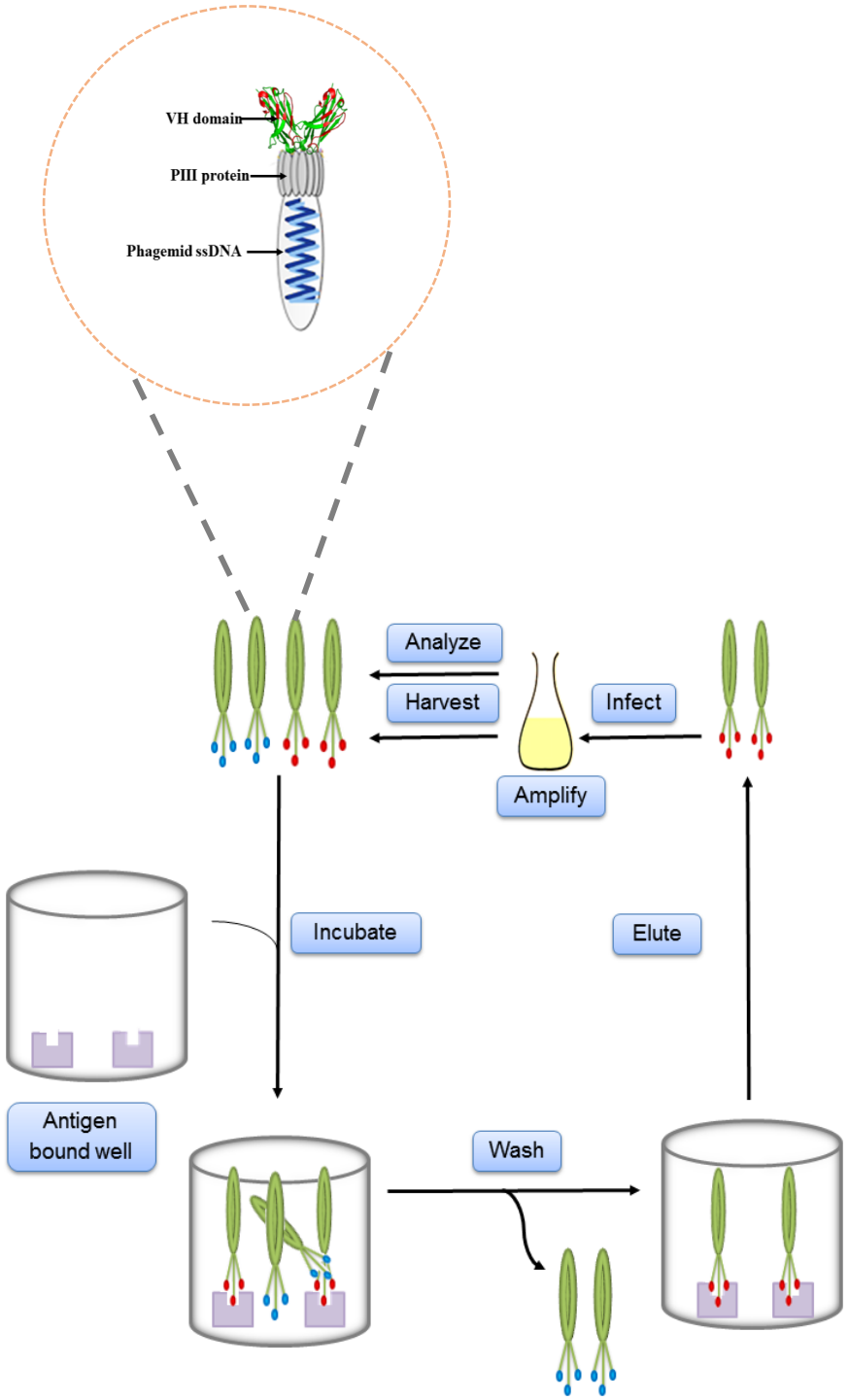


Figure 2.4. Phage Display Selection
(A) Filamentous M13 bacteriophage structure. VHs are fused to the PIII coat protein and displayed on the outside surface of the bacteriophage. (B) Phage displayed VH domain libraries are incubated in a well coated with the target protein. The unbound phages are washed away, and bound phages are eluted and amplified through infection into a bacterial host. Phages are harvested and analyzed for another round of enrichment.

3. OBJECTIVES AND SPECIFIC AIMS

The development of antibody fragments, such as VH domains, circumvents many of the limitations faced with antibodies. However, VH domains have their unique set of limitations such as a high propensity to aggregate. Due to the enormous potential of antibody fragments in diagnostic and therapeutic applications, there is a desire to overcome the inherent instability of these fragments. The objective of this thesis is to increase the thermostability, reversible folding ability, and solubility of human VH domains by altering the charge of the VH domain.

Miklos research group showed that increases in net charge reduced the aggregation of scFvs (Miklos *et al.*, 2012). Perchiacca and co-workers observed that negatively charged mutations near the edges of hydrophobic CDRs inhibit aggregation of negative or near-neutrally charged VH scaffold, and positively charged mutations inhibit aggregation of positively charged VH scaffold at neutral pH (Perchiacca *et al.*, 2014).

Based on these studies we hypothesized that:

1. VH domain thermostability, reversible folding ability, and solubility can be increased by mutating the VH domain framework to increase its net charge.

2. The thermostability, reversible folding ability, and solubility of VH domains can be increased by increasing the charge on CDR1 that can be accommodated by like charged framework region

To achieve our objective the following aims were performed:

3.1. Specific Aim 1: Assess the influence of VH domain framework net charge on thermostability, reversible folding, and solubility.

3.2. Specific Aim 2: Assess the ability of VH domain framework net charge on tolerating charge in CDR1.

4. MATERIALS AND METHODS

4.1. General information

4.1.1. Reagents and Suppliers

The reagents used for the experiments in this thesis were all molecular biology or reagent grade. The suppliers and their addresses are shown in Table 4.1. Other special reagents used in this study are listed in Table 4.2. Enzymes are shown in Table 4.3 and specialized lab equipment used in this study are listed in Table 4.4. Oligonucleotides are listed in Table 4.5.

Table 4.1. Names and addresses of suppliers

Supplier	Address
Applied Photophysics Ltd	Leatherhead, UK
Bio Basic Inc.	Markham, Ontario, Canada
Bio-Rad	Mississauga, Ontario, Canada
EMD millipore	Etobicoke, Ontario, Canada
Fisher Scientific	Ottawa, Ontario, Canada
Integrated DNA Technologies (IDT)	Coralville, Iowa, USA
Invitrogen Life Technologies	Burlington, Ontario, Canada
New England Biolabs (NEB)	Ipswich, MA, USA
Qiagen	Mississauga, Ontario, Canada
Sigma-Aldrich	Oakville, Ontario, Canada
Thermo scientific Fisher	Waltham, MA, USA

Table 4.2. Reagents

Reagent	Supplier
Amicon ultra-4 centrifugal filter device	EMD Millipore
BCA protein assay Kit	Thermo scientific Fisher
Gel Purification Kit	Qiagen
His resin	Qiagen
Instant TB media	Fisher Scientific
MicroAmp® Fast Optical 96-Well Plate, 0.1 ml	Fisher Scientific
Oligonucleotides	Integrated DNA Technologies (IDT)
PCR Purification Kit	Bio basic Inc.
Plasmid DNA extraction Kit	Bio basic Inc.
QIAPrep Spin M13 Kit	Qiagen

Table 4.3. Enzymes	
Enzyme	Supplier
*BamH*1	New England Biolabs
HotStar *Taq* polymerase	Invitrogen Life Technologies
*Not*1	New England Biolabs
Phusion High Fidelity *Taq* DNA Polymerase	Invitrogen Life Technologies
T4 DNA ligase	New England Biolabs
T4 Polynucleotide Kinase (PNK)	New England Biolabs
T7 DNA Polymerase	New England Biolabs

Table 4.4. Lab equipment	
Equipment	Supplier
Chirascan^{TM}-plus CD Spectrometer	Applied Photophysics
Micro-Pulser Electroporator	Bio-Rad
Mini-PROTEAN® 3 Cell	Bio-Rad
NanoDrop 2000c spectrophotometer	Thermo Scientific
UV light transilluminator	Bio-Rad

Table 4.5. Oligonucleotides

Oligonucleotides were ordered from Integrated DNA Technologies. Their sequences in 5′ to 3′ orientation are given in the following table. Serial numbers are used to refer the oligonucleotides in the text. For, rev, seq, and Kunkel represent forward primer, reverse primer, sequencing primer, and Kunkel primer, respectively. All forward, reverse, and sequencing primers were suspended in ddH2O at 200 μM concentration. Kunkel and library oligos were suspended in ddH2O at 1 μg/μl concentration.

S. No	Name	Primer Sequence (5′ → 3′)
1	VH-B1a	GAGGTTCAGCTGGTGGAGTCTGGCGGTGGCCTGGTG CAGCCAGGGGGCTCACTCCGTTTGTCCTGTGCAGCTT CTGGCTTCAACATTAAAGCGGCCGCAGGCTGGGTGC GTCGTGCCCCGGGTAAGGGCGAGGAATGGGTTGCAA GCGCGGCCGCACGTTTCACTATAAGCGCAGACACAT CCAAAAACACAGCCTACCTACAAATGAACAGCTTAA GAGCTGAGGACACTGCCGTCTATTATTGTGCCCGCG CGGCCGCATGGGGTCAAGGAACACTAGTCACCGTCT CCTCGCACCATCACCATCACCAT
2	pET-VH for	GCGATGGCCATGGATATCGGAATTAATTCGGAGGTT CAGCTGGTGGAGTCT
3	pET-VH rev	CGCAAGCTTGTCGACGGAGCTCGAATTCGGCTAATG GTGATGGTGATGGTGCG
4	pET T7 promoter seq	TAATACGACTCACTATAGGG

5	Kunkel negV5D, L11E, Q13D, S21E, A23D	GCCAGAAGCGTCACACTCCAAACGGAGTGAGCCCCC TGGGTCCACCTCGCCACCGCCAGACTCGTCCAGCTG AACCTC
6	Kunkel neg S70D	TTTGGATGTGTCTGCGTCTATAGTGAAACGGCC
7	Kunkel pos S21K, A23K	GCCAGAAGCCTTACACTTCAAACGGAGTGAGCCCCC TGG
8	Kunkel pos A40R	CCATTCCTCGCCCTTACCCGGACGACGACGCAC
9	Kunkel pos T68K	TGTGTCTGCGCTTATCTTGAAACGGCCCTTGAC
10	Kunkel CDR1-DDY	GCAGCTTCTGGCTTCAACATTAAAGACGACTATATA GGCTGGGTGCGTCG
11	Kunkel CDR1-DTD	GCAGCTTCTGGCTTCAACATTAAAGACACCGACATA GGCTGGGTGCGTCG
12	Kunkel CDR1-DDD	GCAGCTTCTGGCTTCAACATTAAAGACGACGACATA GGCTGGGTGCGTCG
13	Kunkel CDR1-KTY	CCAGCCTATATAGGTCTTTTTAATGTTGAA
14	Kunkel CDR1-DKY	CCAGCCTATATACTTGTCTTTAATGTTGAA
15	Kunkel CDR1-DTK	CCAGCCTATCTTGGTGTCTTTAATGTTGAA
16	Kunkel CDR1-KKY	CCAGCCTATATATTTCTTTTTAATGTTGAA
17	Kunkel CDR1-DKK	CCAGCCTATCTTTTTGTCTTTAATGTTGAA
18	Kunkel CDR1-KTK	CCAGCCTATCTTGGTCTTTTTAATGTTGAA
19	Kunkel CDR1-KKK	CCAGCCTATCTTTTTCTTTTTAATGTTGAA

4.1.2. Strains

Table 4.6. *E. coli* strains and genotypes

Strain	Genotype	Reference
BL21 (DE3)	F– *omp*T *hsd*SB (rB^-, mB^-) *gal dcm* (*DE3[lacI lacUV5-T7 gene 1 ind1 sam7 nin5]*)	Novagen
CJ236	FΔ*(HindIII)::cat* (Tra^+ Pil^+ Cam^R)/ *ung-1 relA1 dut-1 thi-1 spoT1 mcrA*	NEB
DH10B	F^-*mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139Δ(*ara*, *leu*)7697 *gal*U *gal*K λ^-*rps*L *nup*G	Fisher Scientific

4.1.3. Plasmids

The following pET-22b(+)/VH plasmid map was created using the free online web source PlasMapper (http://wishart.biology.ualberta.ca/PlasMapper/) (Dong *et al.*, 2004).

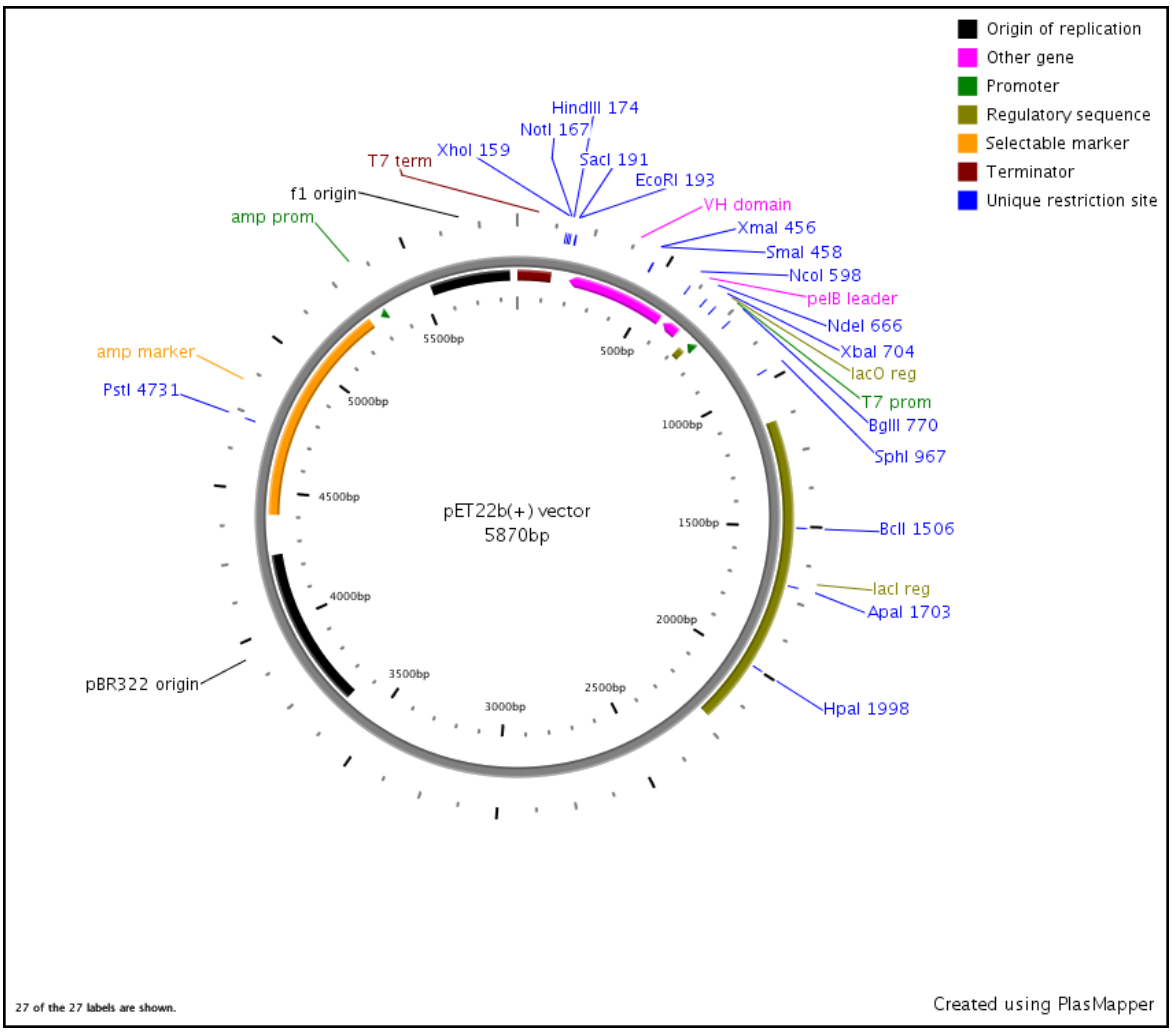


Figure 4.1. pET-22b(+)/VH plasmid
pET-22b(+) was used to express proteins in *E. coli*. VH genes were cloned into *BamH*1 restriction site of pET22b(+) vector. It has an N-terminal pelB signal sequence for potential periplasmic localization of expressed VH domain and C-terminal His-Tag for purification. Expression and termination were controlled by the IPTG-inducible T7 promoter and the T7 terminator sequence, respectively. The F1 origin of replication (F1 ori) was used for single-stranded DNA (ssDNA) production. The pBR322 origin was used to maintain a high copy number of the plasmid in *E. coli*. The ampicillin resistance gene (amp marker) was used to select the plasmid in *E. coli*.

4.2. General protocols

4.2.1. Polymerase Chain Reactions

4.2.1.1. High fidelity PCR

High fidelity PCR reactions were used for VH gene amplification.

Reagents	Amount	Final Conc.
5x Phusion HF buffer (with $MgCl_2$)	10 µL	1x
10 mM dNTPs	1 µL	200 µM each
10 µM forward primer	2.5 µL	0.5 µM
10 µM reverse primer	2.5 µL	0.5 µM
Phusion HF DNA polymerase	0.5 µL	0.02 U/µl
Nuclease-free water	Up to 50 µL	
Template	Variable	50-200 ng
Total volume	50 µL	

PCR products were amplified using the following program in a thermal cycler

	Steps	Temperature	Time
	Initial denaturation	98°C	30 sec
25 Cycles	Denaturation	98°C	10 sec
	Annealing	55°C	30 sec
	Extension	72°C	15-30 sec/kb
	Final extension	72°C	10 min
	Hold	4°C	

4.2.1.2. Low Fidelity PCR

Low fidelity PCR was used to check plasmids for the correct insert.

Reagents	Amount	Final Conc.
10x PCR buffer (with $MgCl_2$)	5 µL	1x
10 mM dNTPs	1 µL	200 µM each
10 µM forward primer	2.5 µL	0.5 µM
10 µM reverse primer	2.5 µL	0.5 µM
HotStart *Taq* DNA polymerase	0.25 µL	2.5 U/reaction
Nuclease-free water	Up to 50 µL	
Template	Variable	50-200 ng
Total volume	50 µL	

PCR products were amplified using the following program in a thermal cycler

	Steps	Temperature	Time
	Initial denaturation	95°C	15 min
25 Cycles	Denaturation	94°C	30 sec
	Annealing	55°C	30 sec
	Extension	72°C	1 min/kb
	Final extension	72°C	10 min
	Hold	4°C	

4.2.2. Agarose gel electrophoresis

PCR products and plasmids were visualized using agarose gel electrophoresis. Samples were mixed with 6x loading dye (50% (v/v) glycerol, 0.2 M EDTA pH 8.3, 0.05% (w/v) bromophenol blue) for a final dye concentration of 1x. Samples were resolved in an agarose gel consisting of 0.8 to 1% (w/v) ultrapure agarose in 1x TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.0) and 0.5 µg/mL ethidium bromide. Electrophoresis was carried out at 100 V for 30 to 60 minutes in 1x TAE buffer and photographed using a UV light transilluminator (Bio-Rad).

4.2.3. Purification and extraction of DNA

PCR products were purified using a PCR Purification Kit (Bio Basic) according to manufacturer's instructions. Restriction enzyme-digested plasmids were purified using Gel Purification Kit (Qiagen) according to manufacturer's instructions. Similarly, gel purification was performed using Gel Purification Kit as per manufacturer's instructions.

4.2.4. DNA sequencing

For the sequencing of all constructed pET22b(+)/VH plasmids, we used sequencing primer listed in 4 of Table 4.5 at a concentration of 5 µM. Plasmid miniprep was used as a sequencing template at a concentration of 50 µg/mL.

4.2.5. Statistical analysis

The Standard Error of Estimate (SEE) for Melting temperature (Tm) was calculated using non-linear regression analysis with SigmaPlot 10.0, where standard Error of Estimate, (SEE) = $\sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{n}}$.

The standard deviation of slope and intercept of Van't Hoff plot and direct fit approach were calculated using LINEST function of Excel. The standard deviation of slope, $s_m = \sqrt{\frac{s^2_{y,x}}{SS_{xx}}}$, where $S^2_{y,x}$ is the variance of y(x) = $\frac{1}{(n-2)}\sum_{i=1}^{n}(y_i - \hat{y}_i)$ and the sum of square of x, SS_{xx}= $\sum_{i=1}^{n}(x_i - \bar{x})^2$. Standard deviation of intercept, sb = $\sqrt{\left[s^2_{y,x}\left(\frac{1}{n} + \frac{\bar{x}^2}{SS_{xx}}\right)\right]}$ (Morrison, 2014). Here x_i and y_i are the actual score, $\hat{y}_i$ is the predicted score and $\bar{x}$ is the mean of the data.

P-values for Tm and fraction folded (α) were calculated from curve fitting with a Boltzmann sigmoidal curve, and compared with Prism 7.0 (Graphpad). P-Values for temperature induced aggregation and room temperature solubility were calculated from multiple comparisons using ordinary one-way ANOVA Prism 7.0 (GraphPad).

4.3. General *E. coli* protocols

4.3.1. Bacterial media

2x Yeast Extract and Tryptone Broth (2YT): 2YT media was prepared with 1.6% (w/v) tryptone, 1.0% (w/v) yeast extract, 85.6 mM NaCl, and 1 mM NaOH in ddH2O. Solid media contained 1.5% (w/v) agar.

Lysogeny broth (LB): LB broth was prepared with 1% tryptone, 0.5% yeast extract, 85.6 mM NaCl, and 1 mM NaOH in ddH2O. The appropriate antibiotic was added after the media cooled to 55°C. Solid media contained 2% agar.

Overnight express instant TB medium: Instant TB media was prepared with 60 g media and 10 mL glycerol in 1 L of ddH2O.

Super Optimal Broth with Catabolic repressor Medium (SOC): SOC media was prepared with 2% (w/v) peptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM Glucose.

Antibiotics: Antibiotics were prepared at a 1000x stock in ddH_2O (or ethanol for chloramphenicol) and stored at -20°C. The appropriate antibiotics were added at concentrations listed in Table 4.7 after the media had cooled to 55°C.

Table 4.7. Antibiotics

Antibiotic	Working conc. (μg/mL)	Supplier
Carbenicillin	50	Sigma-Aldrich
Kanamycin	25	Sigma-Aldrich
Chloramphenicol	10	Sigma-Aldrich

4.3.2. Strain propagation

Standard techniques were used to culture and propagate *E. coli* (Elbing and Brent, 2001). Unless otherwise noted, liquid cultures were grown at 37°C with shaking at 200 rpm. Cultures on solid media were incubated for overnight at 37°C.

4.3.3. Plasmid DNA preparation

Plasmid DNA was prepared by inoculating 5-10 mL of media, containing the appropriate antibiotic with a single colony and grown to saturation. Cells were collected by centrifugation at 4,000 × g for 5 minutes, and the supernatant was removed. Plasmid DNA was obtained using Bio basic Plasmid DNA extraction Kit as described by the manufacturer's instructions. DNA concentration was determined using the NanoDrop 2000c spectrophotometer (Thermo Scientific).

4.3.4. Preparation of electrocompetent *E. coli*

Bacterial strain was streaked on an appropriate agar plate to obtain single bacterial colonies. A single bacteria colony was inoculated into 5 mL of LB and grown overnight at 37°C in a shaking incubator. The overnight culture was diluted in 500 mL of LB medium to an OD_{600} = 0.6. All successive steps were performed at 4°C using pre-chilled tubes. The cells were centrifuged at 4,000 × g for 20 minutes at 4°C and washed twice with 250 mL of ddH2O. The cells were resuspended in 20 mL of ice cold sterile water, transferred to a 50 mL Falcon tube, and centrifuged for 10 minutes at 4,200 × g at 4°C. The pellet was resuspended in 50 mL of

10% glycerol and centrifuged at 4,200 × g at 4°C. The pellet was resuspended in one pellet volume of ice-cold 10% (v/v) glycerol and 50 µL aliquots were stored at -80°C.

4.3.5. *E. coli* transformation

Bacterial cells were transformed by electroporation. One microlitre of plasmid DNA at a concentration of 50-150 ng/µL was mixed with 50 µL of competent cells. The mixture was transferred to an ice-cold electroporation cuvette. Cells were then electroporated using a field strength of 12.5 kV/cm (Ec2 on Bio-Rad Micro Pulser). Electroporated cells were rescued with 500 µL of SOC medium and incubated for 30 minutes at 37°C. Cells were plated onto agar plates containing the appropriate antibiotic, and grown overnight at 37°C.

4.4. Characterization of variable heavy domain (VH)

4.4.1. Variable heavy domain (VH) engineering

VH-B1a gene (listed in 1 of Table 4.5), which was designed by Barthelemy *et al*. (Barthelemy *et al.*, 2008), was ordered from IDT with the incorporation of His-tag at the C-terminal end. VH-B1a was cloned into pET22b(+) vector using Gibson assembly cloning method. The VH-B1a gene was PCR amplified using the primer listed in 2 and 3 of Table 4.5. Amplified VH-B1a was ligated into *BamH*1 restriction site of pET22b(+) vector (Novagen) by following Gibson assembly method (Rouet *et al.*, 2012). Gibson assembled product was transformed into DH10B (Section 4.3.5), and grown for overnight at 37°C with shaking at 200 rpm (Section 4.3.2). Plasmid DNA was extracted following section 4.3.3. Clone of correct sequence was detected through DNA sequencing using primer 4 of Table 4.5 (Section 4.2.4).

Jalview analysis (http://www.jalview.org/) of Miklos designed VH domain of K-neg-2-scFv and K-pos-1-scFv (Miklos *et al.*, 2012) was performed with our designed VH-B1a domain. Six negatively charged mutations from VH domain of K-neg-2-scFv and four positively charged mutations from VH domain of K-pos-1-scFv were chosen. PyMOL analysis (https://www.pymol.org/) was performed to locate these mutations on the VH domain. Getarea tool (http://curie.utmb.edu/getarea.html) was employed to estimate the Solvent Accessible Surface Area (SASA) of individual residues of VH-B1a. The -VH and +VH domain were generated through the introduction of six negative and four positive charge mutations in the

framework region of VH-B1a through Kunkel mutagenesis method (Section 4.5) (Kunkel *et al.*, 1987) using Kunkel primer listed in (5 and 6) and (7-9) of Table 4.5, respectively.

CDR1-mutated VH-B1a and +VH domains were generated through Kunkel mutagenesis method (Section 4.5) (Kunkel *et al.*, 1987)) using Kunkel primer listed in 10-19 of Table 4.5. Purified Kunkel product was transformed into DH10B (Section 4.3.5), and grown for overnight at 37°C with shaking at 200 rpm (Section 4.3.2). Plasmid DNA was extracted following section 4.3.3. Clone of correct sequence was detected through DNA sequencing using primer 4 of Table 4.5 (Section 4.2.4). The isoelectric point (pI), and theoretical net charge of CDR1-mutated VH domains were calculated (http://protcalc.sourceforge.net/) at pH 7.4.

4.4.2. VH expression in autoinduction media

VH constructs in the pET22b(+) vector were transformed into electrocompetent (BL21(DE3)pLysS, Stratagene) cells (Section 4.3.5) and plated on 2YT/carb plates. A single colony was inoculated to 3 ml of 2YT/carb media and grown at 37°C with 200 rpm until OD_{600} = 0.6. Starter cultures were transferred to 200 ml of autoinduction media supplemented with carb and grown for overnight at 30°C with 200 rpm for overnight.

4.4.3. IPTG-based induction of VH expression

Individual colonies were picked from 2YT/carb plates, inoculated into 6 ml 2YT media, and grown at 37°C until OD_{600} = 0.6. Starter cultures were transferred to 200 ml 2YT/carb media and grown for overnight at 30°C with shaking at 200 rpm. Cells were isolated by centrifugation at 16,000 × g for 20 minutes at 4°C and resuspended in 200 mL 2YT/carb media. Bacterial cultures were incubated in a shaking incubator at 30°C for 1 hour before doing induction with 1 mM IPTG (Bio basic). Induction was allowed to proceed for 48 hours at 16°C.

4.4.4. Purification of VH domain

The culture supernatant of VH-B1a, +VH and their CDR1-mutated VH domains were collected by centrifugation at 16,000 × g for 20 minutes at 4°C and incubated overnight with His resin (Qiagen) at 4°C with mild agitation. The cells of -VH domain were lysed, centrifuged, and lysates were collected and incubated with His resin (Qiagen) at 4°C with mild agitation for overnight. The beads were collected, washed, and VH domains were eluted with His elution

buffer with 300 mM imidazole. Finally, buffer exchanged was performed into 1x PBS using the Amicon Ultra-4 centrifugal filter device (EMD Millipore). VHs purity was assessed using SDS-PAGE analysis (12% acrylamide tris-glycine gel, Invitrogen), and concentration was measured using the bicinchoninic acid (BCA) protein assay (Pierce).

4.4.5. Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Purified VHs and bacterial samples were prepared by suspending them in 4x SDS loading dye (100 mM Tris-HCl pH 6.8, 20% (v/v) glycerol, 4% (w/v) SDS, 16% (v/v) 2-mercaptoethanol, 0.2% (w/v) bromophenol blue) for a final concentration of 2x, boiled for 5 minutes and cooled before loading.

SDS-PAGE was performed using a Mini-PROTEAN® 3 Cell (Bio-Rad) as described by Laemmli (Laemmli, 1970). Sample fractions were resolved on a 12% polyacrylamide gel (1.5 M Tris pH 8.8, 10% SDS, 30% acrylamide mix (29.2% acrylamide and 0.8% N, N-methylene-bis-acrylamide), 0.04% (v/v) TEMED, and 10% (w/v) ammonium per sulphate) with a 4% stacking gel (1 M Tris pH 6.8, 10% SDS, 30% acrylamide mix (29.2% acrylamide and 0.8% N, N-methylene-bis-acrylamide), 0.1% (v/v) TEMED, and 10% (w/v) ammonium persulphate) in 1x running buffer (25 mM Tris base pH 8.3, 192 mM glycine, 0.1% (w/v) SDS) at 200 V for 55 minutes.

4.4.6. Coomassie-staining

For Coomassie-staining, SDS-PAGE gels were incubated in a staining solution (0.1% (w/v) Coomassie Brilliant Blue R-250, 50% (v/v) methanol, 10% (v/v) acetic acid) for 1 hour and washed with destaining solution (40% methanol and 10% acetic acid) for 4 hours to overnight. Destained gels were visualized and scanned using UV light transilluminator (Bio-Rad).

4.4.7. Antibody solubility analysis

The temperature-induced aggregation was measured by preparing VH domains at 15 μM (1x PBS, pH 7.4), and heating VH domain samples from 30-90°C for 20 minutes. Samples were cooled at room temperature for 20 minutes and sedimented at 12,000 × g. Long-term

solubility was measured by incubating VH domains at 50 μM (1x PBS, pH 7.4) for seven days at room temperature (25°C). VH domain samples were collected after every 24 hours interval, and aggregates were removed through sedimentation at 10,000 × g. The soluble VH concentrations were assessed using the BCA protein assay (Pierce).

4.4.8. Circular dichroism measurements

Circular dichroism (CD) spectra (http://www.pccf.usask.ca/cdspectrometer.htm) of VH domains were measured at 20 μM concentration (1x PBS, pH 7.4) using Chirascan™-plus CD Spectrometer (Applied Photophysics) (0.5 mm path length, 1 second integration time) in the far-UV region (200-260 nm). Temperature-induced denaturation curve of each VH domain was calculated by monitoring the ellipticity at 207 nm with a heating rate of 5°C/minute.

4.5. Kunkel mutagenesis

4.5.1. Template purification

A single colony of CJ236 from 2YT/carb plate, containing the pET22b(+)/VH plasmid, was used to inoculate 1 mL of 2YT/carb/cmp medium supplemented with M13K07 helper phage and incubated at 37°C with shaking at 200 rpm. After 2 hours, kanamycin was added to select for clones that were co-infected with the M13K07 helper phage. Infected cells were grown with shaking at 37°C for 6 hours and transferred to 30 mL of 2YT/carb/kan/uridine medium. The medium was incubated for 20 hours at 37°C with 200 rpm shaking.

An overnight culture of phage-infected *E. coli* was centrifuged at 4°C for 10 minutes at 16,000 × g. The supernatant was transferred to a tube containing 1/5 volume of PEG/NaCl and incubated for 5 minutes at room temperature. The mixture was centrifuged at 4°C for 10 minutes at 12,000 × g. The supernatant was discarded and centrifuged for an additional 2 minutes to remove residual supernatant. The phage pellet was resuspended in PBS-Tween. To pellet the remaining insoluble matter, the sample was centrifuged at 27,000 × g and 4°C for 5 minutes.

Purification of the uracil-containing ssDNA (dU-ssDNA) was performed at room temperature using a modified version of the Qiagen QIAprep Spin M13 kit protocol. Following phage purification, 7 μL of buffer MP (Qiagen) was added to 500 μL of the phage solution and

incubated for 2 minutes. The sample was added to a QIAprep spin column (Qiagen) and centrifuged for 30 seconds at 10,000 × g. The flow-through was discarded, and the phage particles remain bound to the column matrix. The column was washed twice with 0.7 mL of buffer MLB (Qiagen) and centrifuged 30 seconds at 10,000 × g to separate the DNA from the phage coat protein. The DNA remained bound to the column matrix, and the flow through containing protein coat was discarded. The column was washed by the addition of 0.7 mL of buffer PE (Qiagen), centrifuged 30 seconds at 10,000 × g discarding the flow-through. The column was re-washed with buffer MLB and buffer PE, centrifuged 30 seconds at 10,000 × g to remove residual salts and proteins. Columns were centrifuged in a fresh microcentrifuge tube for an additional 30 seconds to remove any residual PE buffer. The dU-ssDNA was eluted from the column by the addition of 100 µL of buffer EB (Qiagen; 10 mM Tris-HCl, pH 8.5) to the center of the column membrane, incubated for 10 minutes at room temperature, and centrifuged at 10,000 × g for 1 minute into a clean microcentrifuge tube.

4.5.2. Synthesis of Covalently Closed Circular dsDNA (CCC-dsDNA)

First, oligonucleotides listed in 5-19 of Table 4.5 were phosphorylated using 0.6 µg of the mutagenic oligonucleotide, 2.0 µL 10x TM buffer, 2.0 µL 10 mM ATP, 1.0 µL 100 mM DTT, and ddH_2O for a final volume of 20 µL. After the addition of 10 units of T4 polynucleotide kinase, the reaction was incubated for 1 hour at 37°C. After the phosphorylation reaction was complete, oligonucleotides were immediately annealed to the dU-ssDNA template under the following conditions: 1 µg of dU-ssDNA, 2.5 µL 10x TM buffer, 2 µL of each of the phosphorylated oligonucleotides, and ddH_2O to a final volume of 25 µL. The reaction was incubated at 90°C for 3 minutes, 50°C for 5 minutes, and 20°C for 5 minutes. The CCC-dsDNA was synthesized from the annealed oligonucleotide/template mixture by the addition of 1 µL of 10 mM ATP, 1 µL of 25 mM dNTP mix, 1.5 µL of 100 mM DTT, 30 units of T4 DNA ligase, and 30 units of T7 DNA polymerase. The reaction was incubated overnight at 20°C. Next day, DNA was purified and desalted using the PCR Purification Kit (Bio Basic) according to manufacturer's instructions.

5. RESULTS

5.1. Specific Aim 1: Assess the influence of VH domain framework net charge on thermostability, reversible folding, and solubility

5.1.1. Introduction

VH domains are very susceptible to aggregation when they are separated from the VL domain, due to the exposure of hydrophobic amino acid residues at the VH/VL interface to the polar solvent (Barthelemy *et al.*, 2008). By introducing hydrophilic amino acids in the VH/VL interface or by introducing mutations that create a cavity to bury hydrophobic residues from the external environment, it is possible to increase the stability and solubility of the VH domain (Perchiacca and Tessier, 2012). We designed a negatively and positively charged VH domain based on autonomously folding human VH domain (VH-B1a) discovered by Barthelemy *et al.*, (Barthelemy *et al.*, 2008), and analyzed changes in thermostability, reversible folding, and solubility**.**

5.1.2. Design and construction of near-neutrally charged VH domain

We chose a thermodynamically stable and near neutrally charged (+0.3 at pH 7.4) human VH domain (VH-B1a) as a starting point to increase thermostability, solubility, and reversible folding ability of the VH domain. VH-B1a was created by replacing exposed hydrophobic residues of VH-4D5 with five structurally compatible mutations (H35G, Q39R, L45E, R50S, and S93A) (Barthelemy *et al.*, 2008) (Figure 5.1A).

Among these five mutations, Q39R and L45E mutations are oppositely charged and hence stabilize the folded conformation through complementary electrostatic interactions. H35G and R50S mutations are located at the edges of CDR1 and CDR2 and they increase the folding stability by changing the orientation of adjacent aromatic amino acid residues (Barthelemy *et al.*, 2008; Perchiacca and Tessier, 2012). For example, a cavity is produced when histidine at 35-position of VH-4D5 is replaced with glycine, which is occupied by tryptophan at 95-position of CDR3. Thus, sequestration of the hydrophobic tryptophan side chain from the external environment increases the solubility of VH-B1a (Jespers *et al.*, 2004) (Figure 5.1B).

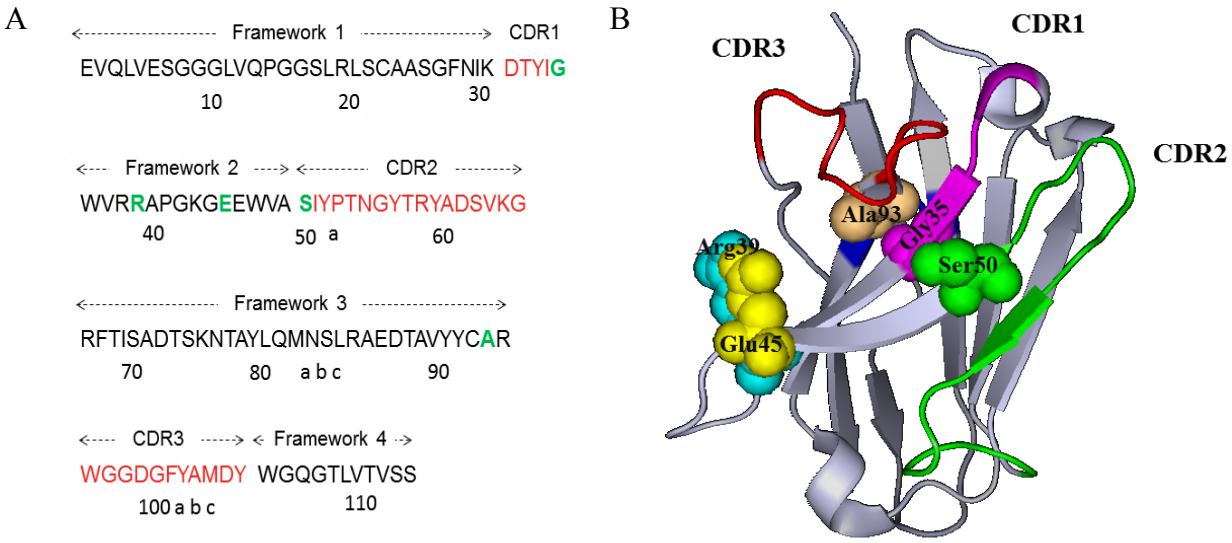


Figure 5.1. Amino acid sequence and crystal structure of VH-B1a
(A) Amino acid sequence of the human VH-B1a domain. CDRs are defined using Kabat numbering and are colored red. Green residues denote amino acids that were mutated in VH-4D5 to produce VH-B1a. (B) Crystal structure of the VH-B1a (PDB: 3B9V). CDR1 is highlighted in pink, CDR2 in green, and CDR3 in red. Gly35 is represented by pink, Arg39 by cyan, Glu45 by yellow, Ser50 by green, and Ala93 by the brown sphere. The blue line shows the disulfide bond between Cys22 and Cys92.

We designed a gene encoding VH-B1a with a His-tag at the C-terminus (S. No 1 of Table 4.5). We PCR-amplified the VH-B1a gene using primers 2 and 3 (Table 4.5) and cloned it into *BamH*1 restriction site of pET22b(+) vector using the Gibson Assembly cloning method (Section 4.4.1) (Rouet *et al.*, 2012).

5.1.3. Design and construction of charged VH domains

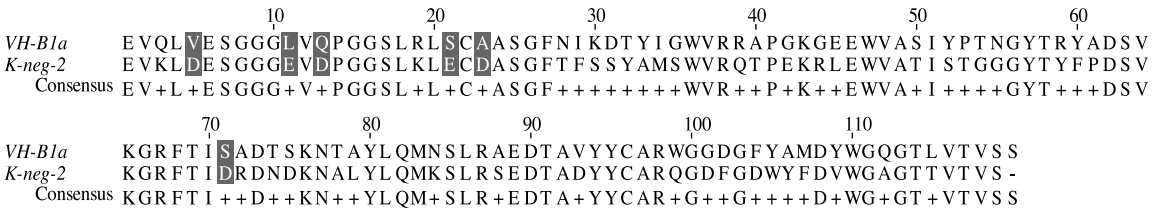
VH domain stability can be improved by increasing the net charge of the framework region without mutating CDRs (Lawrence *et al.*, 2007; Miklos *et al.*, 2012). To observe the effect of net framework charge on the stability and solubility of VH-B1a, we designed a positively and a negatively charged VH domain by introducing positive and negative charged residues in the framework region of the VH-B1a. For this purpose, we followed the anti-MS2 scFv designed by Miklos *et. al.,* where they used a computational package called "Rosetta macromolecular modeling suite" (Das and Baker, 2008) to evaluate the impact of charged mutations on the folding stability of scFvs (Miklos *et al.*, 2012). They selected charged mutations in both VH and VL domains of anti-MS2 scFv, which were anticipated not to affect

its folding. They generated a panel of positively and negatively charged scFv variants by replacing a different number of solvent-exposed charged residues. They observed that the target binding activity of three out of four negatively charged scFvs were significantly reduced after heating, whereas three out of five positively charged scFvs retained significant binding activity.

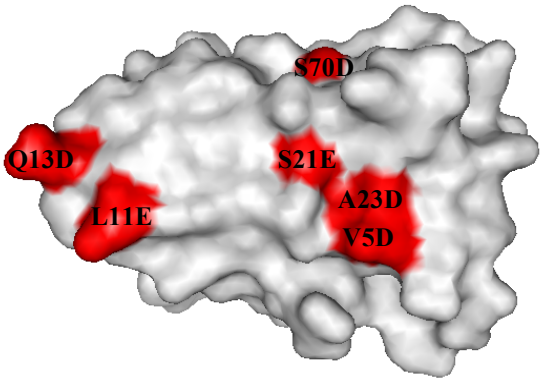
5.1.3.1. Negatively charged VH domain

We designed a negatively charged VH (-VH) domain by introducing negatively charged residues in the framework region of VH-B1a to observe how the negative framework charge affects the stability and solubility of VH-B1a. We designed a -VH domain by following K-neg-2-scFv designed by Miklos *et al*., as it retained strong target binding affinity. We aligned the amino acid sequence of VH-B1a and VH domain of K-neg-2-scFv using Jalview (http://www.jalview.org/) to extrapolate where the negatively charged mutations (V5D, L11E, Q13D, S21E, A23D, and S70D) were located on the VH-B1a (Miklos *et al*., 2012) (Figure 5.2A).

A



B



C

Residue	SASA (angstrom^2)	RSA (%)
VAL-5	84.44	62.90
LEU-11	75.24	51.40
GLN-13	109.65	76.00
SER-21	3.79	Inside
ALA-23	32.02	49.10
SER-70	50.13	62.50

Figure 5.2. Design of -VH domain
(A) Sequence alignment of VH-B1a and VH domain of K-neg-2-scFv. Gray highlighted amino acids represent six negatively charged mutations introduced into VH-B1a. (B) Surface representation of VH-B1a showing that six mutations (Red) are located on the surface of the VH-B1a domain. (C) Solvent Accessible Surface Area (SASA) and relative solvent accessibility (%) of the six mutated amino acids.

We manually analyzed the 3D-structure of VH-B1a through PyMOL (https://www.pymol.org/) and observed that these mutations were located on the exterior surface of VH-B1a (Figure 5.2B). Additionally, we used Getarea tool (http://curie.utmb.edu/getarea.html) to estimate the Solvent Accessible Surface Area (SASA) of individual residues of VH-B1a (Fraczkiewicz, 1998a, 1998b) and found that except for serine at 21-position the other five mutations were solvent exposed (Figure 5.2C). We introduced six framework mutations (V5D, L11E, Q13D, S21E, A23D, and S70D) into VH-B1a to design a -VH domain with a theoretical net charge of -5.7 at pH 7.4 and theoretical pI of 5.69 (http://protcalc.sourceforge.net/).

5.1.3.2. Positively charged VH domain

We designed a positively charged VH (+VH) domain by introducing positively charged residues in the framework region of VH-B1a and examined the effect of net positive framework charge on the stability and solubility of VH-B1a. We designed a +VH domain by following the positive scFv (K-pos-1) designed by Miklos *et al.* that showed high target affinity and thermal resistance. Similar to the -VH domain, we aligned the amino acid sequence of VH-B1a and VH domain of K-pos-1-scFv using Jalview (http://www.jalview.org/) to extrapolate where the positive charged mutation on the VH domain of K-pos-1-scFv aligned with VH-B1a (S21K, A23K, A40R, and T68K) (Miklos *et al*., 2012) (Figure 5.3A).

We manually analyzed the 3D-structure of VH-B1a through PyMOL (https://www.pymol.org/) and observed that these mutations were located on the surface of VH-B1a (Figure 5.3B). Additionally, we used Getarea tool (http://curie.utmb.edu/getarea.html) to estimate the Solvent Accessible Surface Area (SASA) of individual residues of VH-B1a (Fraczkiewicz, 1998a, 1998b) and found that except for serine at 21-position other three mutations were solvent exposed (Figure 5.3C). We introduced four framework mutations (S21K, A23K, A40R, and T68K) into VH-B1a to design a +VH domain with a theoretical net charge of +4.3 at pH 7.4 and theoretical pI of 9.04 (http://protcalc.sourceforge.net/).

To construct -VH and +VH domain, we made dU-ssDNA from VH-B1a cloned in the pET22b(+) vector (Section 4.5.1). We used Kunkel mutagenesis method with VH-B1a dU-ssDNA and mutagenic oligonucleotides 5-9 (Table 4.5 and Section 4.5.2) (Kunkel *et al.*, 1987). Purified Kunkel mutation reaction products were used to transform into *E. coli* DH10B (Section

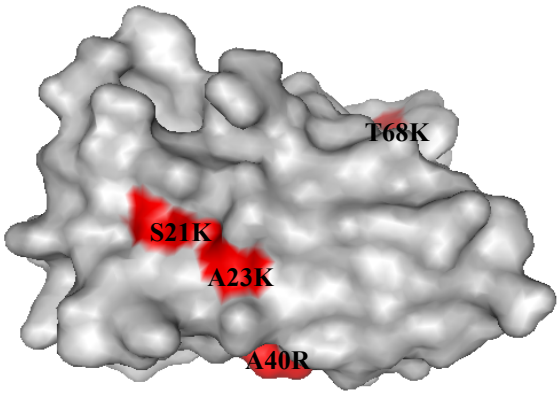
4.3.5). Plasmids were amplified and purified (Section 4.3.3) and following sequence verification (Section 4.2.4) plasmids were used for VH domain expression (Section 4.4.2 and 4.4.3) and purification (Section 4.4.4).

A

```
                    10        20        30        40        50        60
VH-B1a      EVQLVESGGGLVQPGGSLRLSCAASGFNIKDTYIGWVRRAPGKGEEWVASIYPTNGYTRYADSV
K-pos-1     EVKLVESGGGLVKPGGSLKLKCKASGFTFSSYAMSWVRQRPEKRLEWVATISTGGGYTYFPDSV
Consensus   EV+LVESGGGLV+PGGSL+L+C+ASGF++++++++WVR++P+K++EWVA+I++++GYT+++DSV

                    70        80        90        100       110
VH-B1a      KGRFTISADTSKNTAYLQMNSLRAEDTAVYYCARWGGDGFYAMDYWGQGTLVTVSS
K-pos-1     KGRFKISRDNAKNALYLQMKSLRSEDTARYYCARQGDFGDWYFDVWGAGTTVTVS-
Consensus   KGRF+IS+D++KN++YLQM+SLR+EDTA+YYCAR+G++G++++D+WG+GT+VTVSS
```

B



C

Residue	SASA (angstrom^2)	RSA (%)
SER-21	3.79	Inside
ALA-23	32.02	49.10
ALA-40	23.38	29.30
THR-68	76.08	71.10

Figure 5.3. Design of +VH domain
(A) Sequence alignment of VH-B1a and VH domain of K-pos-1-scFv. Gray highlighted amino acids represent four positively charged mutations introduced into VH-B1a. (B) Surface representation of VH-B1a showing that four mutations (Red) are located on the surface of the VH-B1a domain. (C) Solvent Accessible Surface Area (SASA) and relative solvent accessibility (%) of the four mutated amino acids.

5.1.4. Expression optimization of VH-B1a, -VH, and +VH domains

We optimized expression of VH-B1a, -VH, and +VH domains in the BL21 strain under a variety of conditions using autoinduction media and 2YT media with different Isopropyl β-D-1-thiogalactopyranoside (IPTG) concentrations, induction times, and incubation temperatures. We observed that VH-B1a expressed well and was secreted in the culture supernatant both in autoinduction media and 2YT media with 1 mM IPTG induction at 30ºC using a 24 hours incubation period.

We expressed -VH domain both in autoinduction media (Section 4.4.2) and 2YT media with 1 mM IPTG induction (Section 4.4.3), but did not obtain soluble protein in the culture

supernatant. We analyzed the whole cell lysate from the -VH domain expression culture using an SDS gel and observed a band at ~13 kDa (Figure 5.4A).

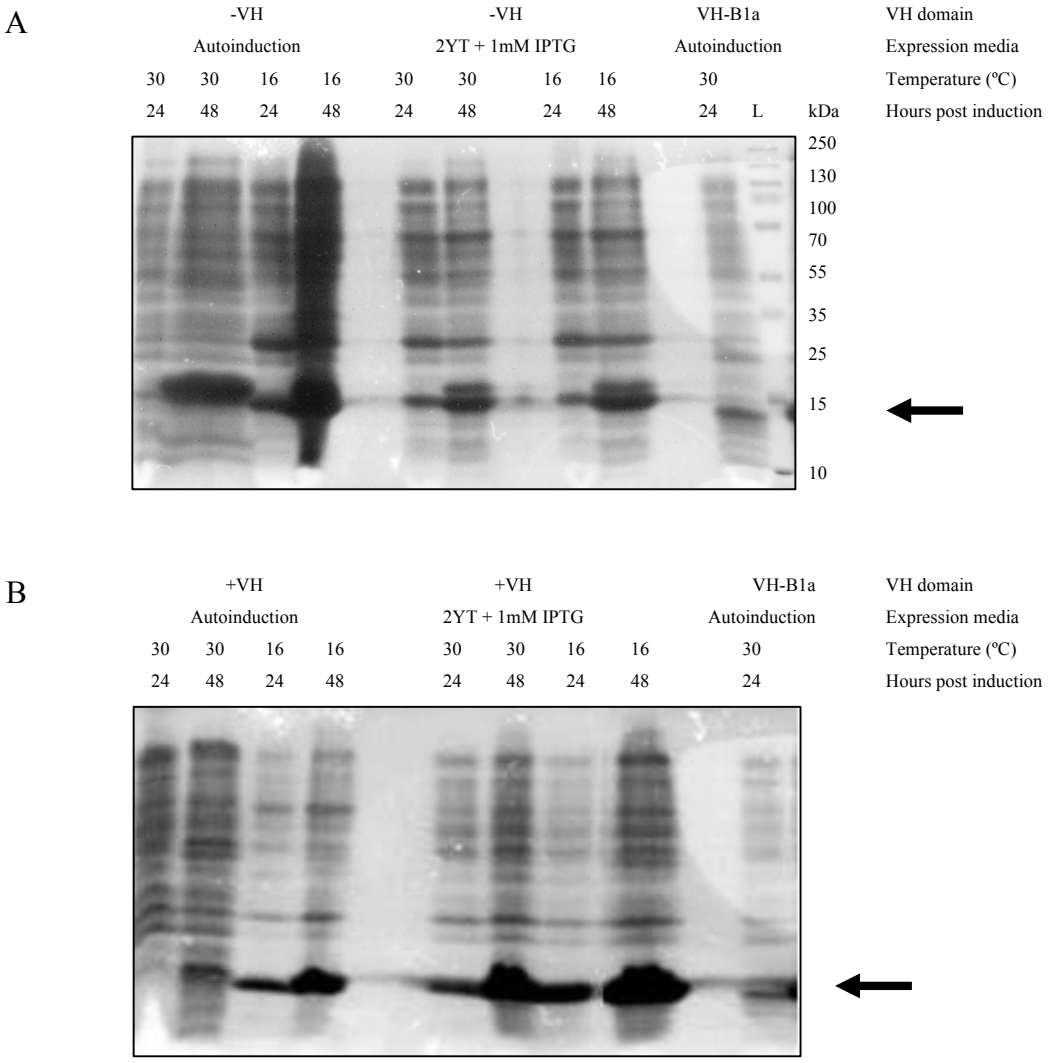


Figure 5.4. Optimization of -VH and +VH expression in the BL21 strain
The -VH (A) and +VH (B) domains were expressed in BL21 *E. coli* using autoinduction media and 2YT media with 1 mM IPTG induction. Following induction, cells were incubated at 16°C and 30°C. Samples were collected 24 hours and 48 hours post-IPTG induction. VH-B1a expression in autoinduction media at 30°C for 24 hours was used to compare -VH and +VH expression. The whole bacterial lysate was resolved on an SDS-PAGE gel. Proteins were visualized by Coomassie staining. Arrows indicate the expected location of the -VH and +VH domain.

We expressed +VH domain in autoinduction media (Section 4.4.2) and did not obtain soluble protein in the culture supernatant. We analyzed the whole cell lysate from the +VH

domain expression culture using an SDS gel and observed a protein band at ~13 kDa position. Next, we expressed +VH domain in 2YT media with 1 mM IPTG induction (Section 4.4.3) and saw an increase in +VH expression in the culture supernatant. However, after IPTG induction, expression of the +VH domain improved at 16°C compared to the expression at 30°C. Further, we observed that 48 hours incubation following induction was sufficient to produce elevated levels of +VH expression (Figure 5.4B).

5.1.5. Purification of VH-B1a, +VH, and -VH domains

After optimization of VH domain expression, we purified VH domains using a His-tag-based purification method (Section 4.4.4). Purified proteins were analyzed using SDS-PAGE (Section 4.4.5) and visualized with Coomassie staining (Section 4.4.6) (Figure 5.5). We observed a lower level of expression for +VH (16 mg/L) and -VH domains (3 mg/L) compared to VH-B1a (20 mg/L).

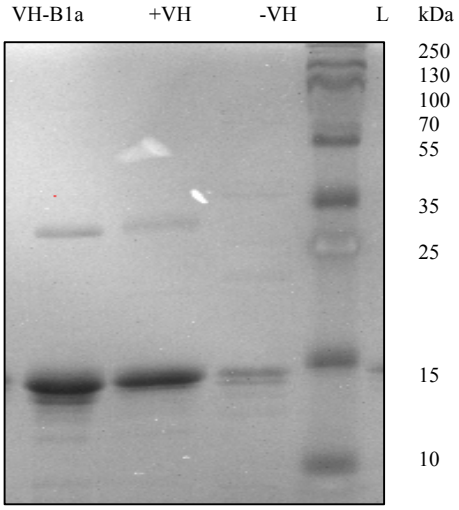


Figure 5.5. SDS-PAGE of purified VH-B1a, +VH and -VH domains
VH-B1a, +VH, and -VH domains were expressed in the BL21 strain using 2YT media with 1 mM IPTG induction and purified by His-tag-based purification method. Purified VH domains were resolved on an SDS-PAGE gel and visualized by Coomassie staining.

5.1.6. Characterization of VH-B1a, +VH, and -VH using circular dichroism

Circular dichroism (CD) is a standard technique to observe the secondary structure, reversible folding ability, and melting temperature (Tm) of proteins by measuring the change in absorption as a function of temperature. As alpha helices and beta sheets are chiral they absorb

circularly polarized light, which is measured by CD spectroscopy. From the CD data, we can calculate thermodynamic properties of proteins (Greenfield, 2006b).

5.1.6.1. Secondary structure analysis of VH-B1a, +VH, and -VH domains

Different protein structural elements have unique CD spectra. For example, α-helical proteins have a negative peak at 222 nm and 208 nm, and a positive peak at 193 nm (Holzwarth and Doty, 1965). Proteins with antiparallel β-pleated sheets (β-helices) have a negative peak at 218 nm and a positive peak at 195 nm (Greenfield and Fasman, 1969). Disordered proteins have very low ellipticity above 210 nm and negative peak near 195 nm (Venyaminov *et al.*, 1993). Thus, CD is widely used to determine the secondary structure of proteins and can be used to monitor the structural changes due to temperature and mutations (Greenfield, 2006b).

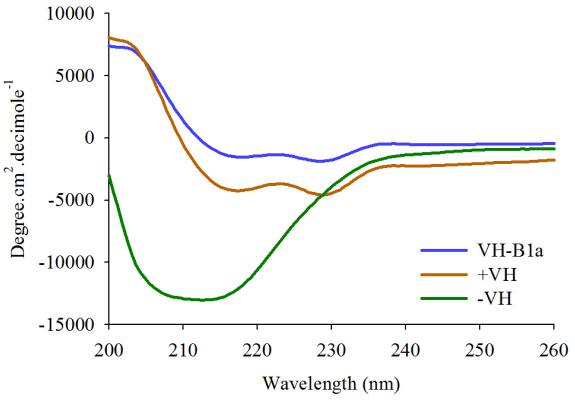


Figure 5.6. Secondary structure of VH-B1a, +VH and -VH domains
Far-UV CD spectra of VH-B1a, +VH, and -VH domains at 30°C. Color lines represent the CD spectra of VH-B1a, +VH and -VH domains at 30°C.

We compared CD data of VH-B1a**,** +VH, and -VH domains in the far-UV (200-260 nm) region at 30°C to observe how the insertion of positively and negatively charged residues in the framework region of VH-B1a affects the secondary structure of +VH and -VH domains. From the CD data, we observed that VH-B1a and +VH domain showed similar secondary structure, whereas -VH domain showed a different secondary structure pattern (Figure 5.6). Four framework mutations (S21K, A23K, A40R, and T68K) of the +VH domain did not substantially perturb the secondary structure, whereas the six framework mutations (V5D, L11E, Q13D, S21E, A23D, and S70D) of -VH domain likely denatured the VH domain. The

CD analysis of the -VH domain indicated that it was not correctly folded and we excluded it from further experiments.

5.1.6.2. Thermostability and reversible folding of VH-B1a and +VH domain

Thermostability and the reversible folding ability of VH-B1a and +VH were calculated by measuring the change in CD spectra (200-260 nm) while heating them from 30°C to 90°C (Figure 5.7).

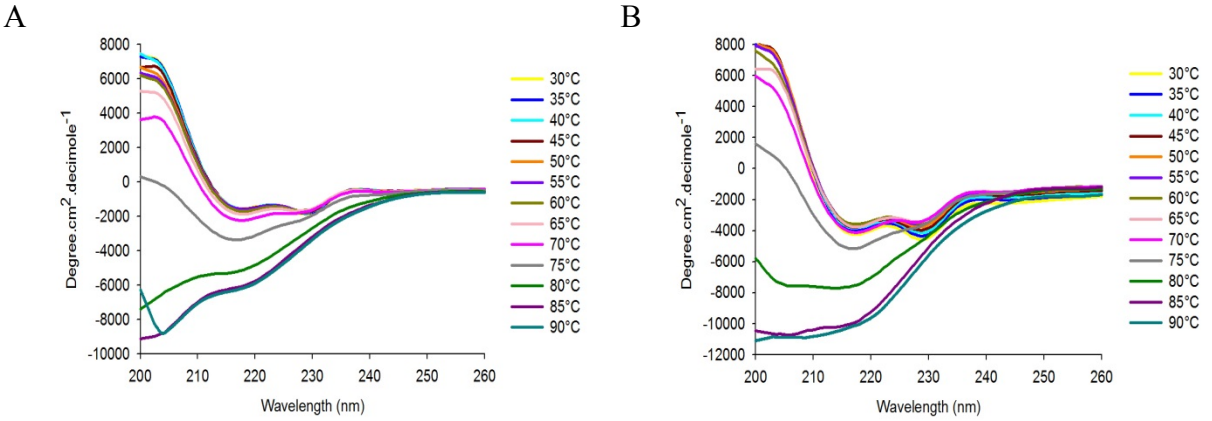


Figure 5.7. CD spectra of VH-B1a and +VH domains
Far-UV CD spectra of VH-B1a (A) and +VH (B) domains collected during first thermal unfolding. Color lines represent the CD spectra of VH-B1a and +VH domains at different heating temperature.

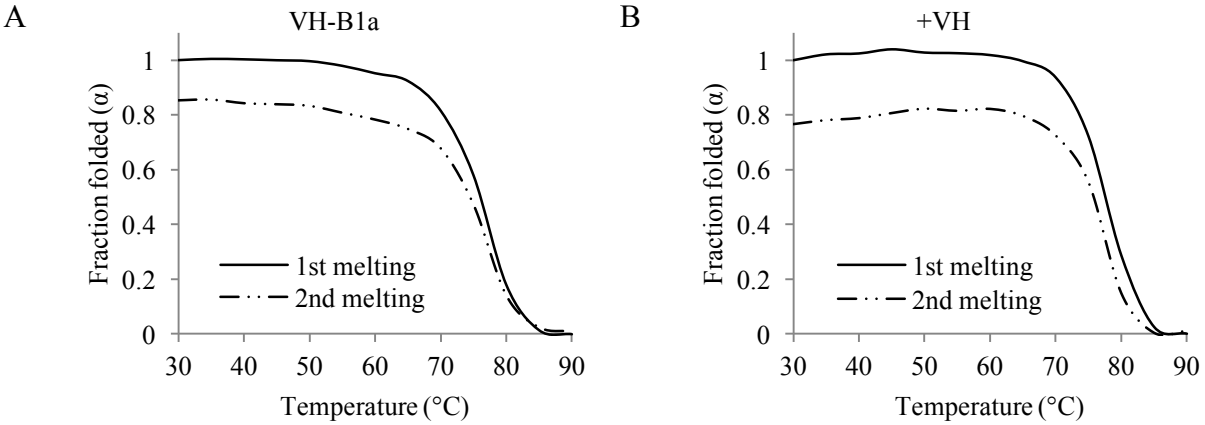
The thermostability of the VH domain was assessed by calculating the melting temperature (Tm) at 207 nm as VH domains are predominantly composed of antiparallel β-pleated sheet structures. At the Tm, the concentration of folded and unfolded protein is equal and equilibrium constant for folding (K_F) = [Folded] / [Unfolded] = 1. The ability of VH domains to unfold reversibly was measured by heating them from 30°C to 90°C (1st melting), cooling them to 30°C, and then reheating them from 30°C to 90°C (2nd melting). The CD spectra during the 1st and 2nd melting were used to calculate the fraction of VH domain folded (α) at a given temperature using the following equation:

$$\alpha = (\Theta_T - \Theta_U) / (\Theta_F - \Theta_U)$$

Where Θ_T is the ellipticity at any temperature, Θ_F is the ellipticity of the fully folded form of VH domain observed at 30°C, and Θ_U is the ellipticity of the unfolded form, which is the

lowest ellipticity value (Greenfield, 2006a). The fraction of VH domain folded (α) for VH-B1a and +VH were 85% and 76%, respectively (Figure: 5.8).

The fraction of VH domain folded (α) was plotted against the temperature to obtain the temperature induced denaturation curve of the VH-B1a and +VH domain. Tm is the temperature at which α is 0.5. The Tm of VH-B1a and +VH were calculated to be 76°C and 78°C, respectively (Figure: 5.8).



Tm = 76 ± 0.03°C
Fraction folded = 85%

Tm = 78 ± 0.03°C
Fraction folded = 76%

Figure 5.8. Temperature-induced denaturation curve of VH-B1a and +VH domains
CD analysis of the first (solid line) and second (dashed line) melting of VH-B1a (A) and +VH (B) domains at 207 nm. Standard Error of Estimate (SEE) was calculated using non-linear regression analysis with SigmaPlot 10.0.

5.1.6.3. Thermodynamics of VH-B1a and +VH domain folding

The thermodynamic parameters calculated from CD data are almost identical to the thermodynamic parameters obtained from other spectroscopic measurements such as fluorescence measurement, if the unfolding of VH domain is reversible, occurs in two states, and shows a single unfolding transition. VH-B1a and +VH domains undergo a reversible folding transition between two states- folded (F) and unfolded (U) when they are exposed to denaturing condition caused by heating. Thus, the thermodynamic parameters of VH unfolding, such as Van't Hoff enthalpy (ΔH), entropy (ΔS), free energy (ΔG_U), equillibrium constant (K_U), and melting temperature (Tm) can be calculated at a particular wavelength by monitoring the change in CD spectra as a function of temperature. Thermodynamics of VH unfolding can be

followed either by collecting the entire spectra as a function of temperature or by monitoring the ellipticity at a particular wavelength (Greenfield, 2006a).

5.1.6.3.1. Thermodynamic parameters of VH unfolding

Thermodynamic parameters of VH unfolding (ΔH, ΔS, ΔG_U, K_U and Tm) can be calculated by analyzing thermal unfolding curves, Van't Hoff's plots, and by a direct fit approach (Pace, 1997; Greenfield, 2006b):

1. Heat-induced thermal unfolding curve of VH domains

The Van't Hoff's and direct fit approach plots were generated using the heat-induced thermal unfolding curve of VH domains (Figure 5.9) (Pace, 1997). The thermal unfolding curve representing ellipticity versus temperature, the change in ellipticity was linear as a function of temperature preceding and succeeding the unfolding transition (Greenfield, 2006a). Thus, we performed a linear least square analysis at the pre- and post-transition regions of VH-B1a and +VH domains and acquired a straight-line equation for Θ_F and Θ_U. From the obtained Θ_F and Θ_U values of the thermal unfolding curve of VH-B1a and +VH domains, we calculated K_U and ΔG_U at the transition region (70°C, 75°C, and 80°C) where the VH-B1a and +VH domains become unfolded (Table 5.1 and 5.2).

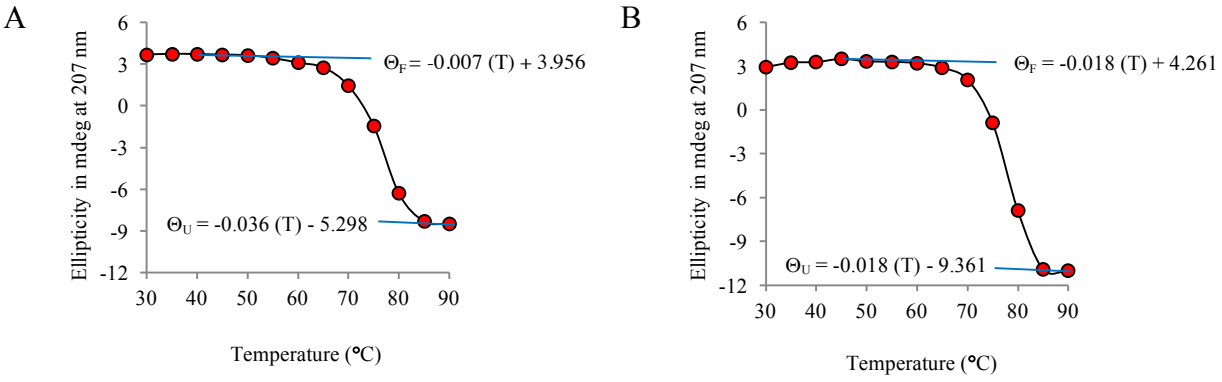


Figure 5.9. Thermal unfolding of VH-B1a and +VH domains measured by CD at 207
Thermal unfolding of VH-B1a (A) and +VH (B) domains were monitored by CD at 207 nm during their first melting from 30°C to 90°C. Θ_F and Θ_U represent the ellipticity of the folded and unfolded states, respectively.

Table 5.1. Analysis of thermal unfolding curve of VH-B1a					
Temp (°C)	Ellipticity in mdeg at 207 nm	[a]Fraction unfolded, (F_U)	[b]Equilibrium constant for unfolding, (K_U)	lnK_U	[c]ΔG for unfolding, (ΔG_U) (cal/mol)
70	1.42	0.18	0.23	-1.49	1013.53
75	-1.45	0.42	0.74	-0.31	211.19
80	-6.30	0.83	4.92	1.59	-1118.01

Ellipticity of the fully folded form, $\Theta_F = -0.007\ (T) + 3.956$
Ellipticity of the fully unfolded form, $\Theta_U = -0.036\ (T) - 5.298$
[a]Fraction unfolded, $(F_U) = (\Theta_F - \Theta_T) / (\Theta_F - \Theta_U)$
[b]Equilibrium constant of unfolding, $(K_U) = (\Theta_F - \Theta_T) / (\Theta_T - \Theta_U)$
[c]Free energy of unfolding, $\Delta G_U = -RTlnK_U$ (cal/mol)

Table 5.2. Analysis of thermal unfolding curve of +VH					
Temp (°C)	Ellipticity in mdeg at 207 nm	[a]Fraction unfolded, (F_U)	[b]Equilibrium constant for unfolding, (K_U)	lnK_U	[c]ΔG for unfolding, (ΔG_U) (cal/mol)
70	2.06	0.08	0.09	-2.45	1669.88
75	-0.91	0.29	0.41	-0.90	618.56
80	-6.90	0.72	2.51	0.92	-646.44

Ellipticity of the fully folded form, $\Theta_F = -0.018\ (T) + 4.261$
Ellipticity of the fully unfolded form, $\Theta_U = -0.018\ (T) - 9.361$
[a]Fraction unfolded, $(F_U) = (\Theta_F - \Theta_T) / (\Theta_F - \Theta_U)$
[b]Equilibrium constant of unfolding, $(K_U) = (\Theta_F - \Theta_T) / (\Theta_T - \Theta_U)$
[c]Free energy of unfolding, $\Delta G_U = -RTlnK_U$ (cal/mol)

2. Calculating thermodynamic parameters of VH domain unfolding using Van't Hoff plot

Van't Hoff equation is used to determine the change in enthalpy (ΔH) and entropy (ΔS) during the thermal unfolding of VH domains. The Van't Hoff equation is derived from Gibbs Free energy equation (Equation 1).

$$\Delta G = \Delta H - T\Delta S \ldots\ldots. (1)$$

ΔH is the enthalpy change and Δ*S* is the entropy change of the system. The Gibbs free energy of unfolding is related to the temperature and equilibrium constant (K_U) (Equation 2).

$$\Delta G = -RTlnK_U \ldots\ldots (2)$$

Substituting equation 1 into equation 2 generates equation 3, which is the linear form of the Van't Hoff equation, where lnK_U is a linear function of 1/T.

$$\ln K_U = -\Delta H/RT + \Delta S/R \ldots\ldots. (3)$$

Since the temperature-induced unfolding of VH-B1a and +VH domains were a reversible reaction, we measured the equilibrium constant of unfolding (K_U) at the transition region, where VH domain become unfolded at 70°C, 75°C, and 80°C (Table 5.1 and 5.2). We generated a Van't Hoff plot (Figure 5.10) to calculate the enthalpy and entropy of VH domain unfolding (Pace, 1997). In the Van't Hoff plot, -ΔH/R is the slope and ΔS/R is the X intercept. Change in enthalpy can be positive or negative, leading to two major forms of the Van't Hoff plot: exothermic and endothermic plot. In an endothermic reaction, heat is absorbed, net enthalpy change is positive (ΔH > 0), and the slope is negative. Moreover, heat is released in an exothermic reaction making the net enthalpy change negative (ΔH < 0) and slope positive. We observed that the temperature-induced unfolding of VH-B1a and +VH domains were endothermic reaction (Figure 5.10).

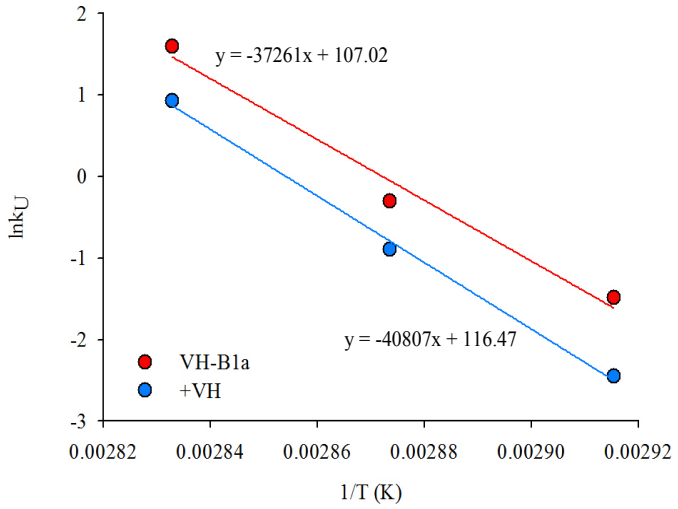


Figure 5.10. Van't Hoff plot of VH-B1a and +VH domains
Van't Hoff plot of $\ln K_U$ versus 1/T yields a straight line with a slope -ΔH/R and X-intercept ΔS/R. The slope of Van't Hoff plot of VH-B1a gives ΔH = -R × (slope ± std dev) = 74.04 ± 10.58 kcal/mol and Y-intercept contributes ΔS = R × (intercept ± std dev) = 212.65 ± 30.42 cal/mol/K. The slope of Van't Hoff plot of +VH gives ΔH = 81.08 ± 4.29 kcal/mol and Y-intercept provides ΔS = 231.43 ± 12.34 cal/mol/K. Red line represents VH-B1a and the blue line represents +VH. The standard deviation of slope and intercept were calculated using LINEST function of Excel.

3. Calculating thermodynamic parameters of VH domain unfolding using Direct fit approach

The simplest method to determine Tm and ΔHm is to plot ΔG as a function of temperature (Table 5.1 and 5.2), referred to as the direct fit approach plot (Figure 5.11). Direct Fit approach plot is derived from the Gibbs Free energy equation (Equation 4):

$$\Delta G = \Delta H - T\Delta S \ldots\ldots. (4)$$

The temperature at which ΔG = 0 and K_{eq} = 1 indicates the melting point (Tm), the midpoint of the thermal unfolding curve at which 50% of the VH domains are folded and 50% are unfolded. At the Tm, equation 5 gives:

$$\Delta G(Tm) = -RTlnK_{eq} \ldots\ldots. (5)$$

$$\Delta G(Tm) = 0 \ldots\ldots. (6)$$

Using equation 4 and equation 6, we get following:

$$\Delta G(Tm) = \Delta Hm - Tm\Delta Sm = 0$$

$$\Delta Hm = \Delta SmTm \ldots\ldots. (7)$$

The slope of the plot is ΔSm and slope multiplied by Tm give ΔHm (Equation 7) (Pace, 1997).

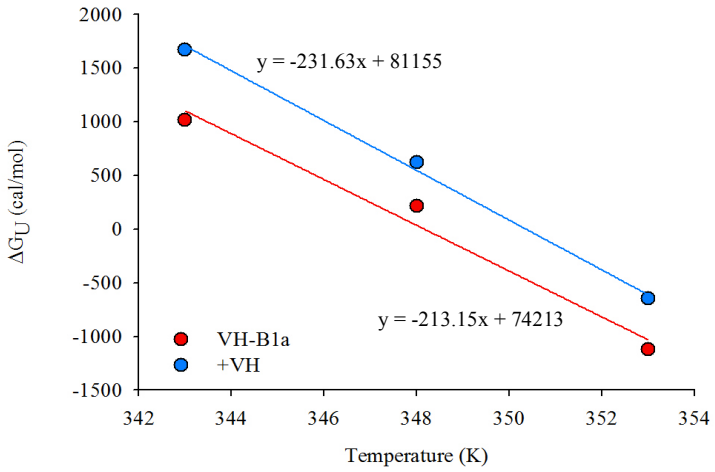


Figure 5.11. Direct Fit approach plot of VH-B1a and +VH domains
The direct fit approach plot of ΔG_U versus Temperature (K), yields a straight line with a slope ΔSm, Tm = T at which ΔG = 0 and ΔHm = TmΔSm. The slope of the VH-B1a gives ΔSm = -(slope ± std dev) = 213.15 ± 30.42 cal/mol/K, Tm = 348 K = 75°C and ΔHm = (Tm) (ΔSm ± std dev) = 74.18 ± 10.59 kcal/mol. The slope of +VH gives ΔSm = 231.63 ± 12.34 cal/mol/K, Tm = 350.50 K = 77.50°C and ΔHm = 81.19 ± 4.29 kcal/mol. Red line represents VH-B1a and the blue line represents +VH. The standard deviation of slope and intercept were calculated using LINEST function of Excel.

The Tm calculated using direct fit approach was similar to the Tm values obtained from temperature-induced denaturation curve (Table 5.5). We obtained nearly identical ΔH and ΔS values calculated from the Van't Hoff plot and Direct Fit approach (Table 5.5). We calculated the change in enthalpy (ΔΔH) and entropy (ΔΔS) for +VH domain relative to VH-B1a and determined whether the positively charged mutations have any effect on unfolding mechanism. The VH domain enthalpy change of +VH domain: ΔΔH = ΔH(VH-B1a) - ΔH(+VH) kcal/mol = ((74.18 ± 10.59) - (81.19 ± 4.29)) kcal/mol = -7.01 ± 14.88 kcal/mol. The VH domain entropy change of +VH domain: ΔΔS = ΔS(VH-B1a) - ΔS(+VH) cal/mol = ((213.15 ± 30.42) - (231.63 ± 12.34)) cal/mol/K = -18.48 ± 42.76 cal/mol/K. From the calculated enthalpy and entropy change, we can conclude that four framework mutations of VH-B1a domain did not change the unfolding mechanism of the +VH domain (Table 5.5).

Table 5.3. Thermodynamic parameters of VH-B1a and +VH domain unfolding

VH domains	ΔG_U[a] kcal/mol	$\Delta\Delta G$[b] kcal/mol	K_U[a]	Tm (°C) ± SEE[d]	Direct fit[e]			Van't Hoff plot[f]		ΔΔH ± std dev[g] kcal/mol	ΔΔS ± std dev[h] cal/mol/K
					Tm (°C)	ΔSm ± std dev (cal/mol/K)	ΔHm ± std dev (kcal/mol)	ΔH_U ± std dev (kcal/mol)	ΔS_U ± std dev (cal/mol/K)		
VH-B1a	9.60 ± 19.81	N/A	1.19 x 10^{-7}	76.00 ± 0.03	75.00	213.15 ± 30.42	74.18 ± 10.59	74.04 ± 10.58	212.65 ± 30.42	N/A	N/A
+VH	11.00 ± 8.03	-1.4 ± 27.84	1.16 x 10^{-8}	78.00 ± 0.03	77.50	231.63 ± 12.34	81.19 ± 4.29	81.08 ± 4.29	231.43 ± 12.34	-7.01 ± 14.88	-18.48 ± 42.76

[a]Free energy of unfolding, $\Delta G_U = \Delta H_U - T\Delta S_U$ kcal/mol
[b]$\Delta\Delta G = \Delta G_U$(VH-B1a) - ΔG_U(+VH) kcal/mol
[c]Equillibrium constant of unfolding, $K_U = e^{-\Delta GU/RT}$
[d]Midpoint of thermal unfolding curve in °C ± SEE
[e]Direct Fit approach: Tm in °C at which ΔG = 0,
ΔSm = - (slope ± std dev) cal/mol/K,
ΔHm = (Tm)(ΔSM ± std dev) kcal/mol
[f]Van't Hoff plot: $\Delta H_U = -R \times$ (slope ± std dev) kcal/mol
$\Delta S_U = R \times$ (intercept ± std dev) cal/mol/K
[g]ΔΔH = ΔH(VH-B1a) - ΔH(+VH) kcal/mol
[h]ΔΔS = ΔS(VH-B1a) - ΔS(+VH) cal/mol/K

5.1.6.3.2. Equilibrium constant and free energy of VH unfolding

The free energy of unfolding and realted equilibrium constant provides information on the favored direction of VH folding and unfolding (Pace, 1997; Greenfield, 2006a). The free energy of unfolding (ΔG_U) was calculated from the enthalpy and entropy of VH unfolding using Gibbs Free energy equation, $\Delta G_U = \Delta H_U - T\Delta S_U$, which represents the stability of VH domain. The equilibrium constant (K_U) was calculated from the free energy of VH unfolding (ΔG_U), $K_U = e^{-\Delta GU/RT}$. Here the ideal gas constant is (R) = 1.987 cal/mol/K and absolute temperature of unfolding is (T) = (273+30) K = 303 K.

Table 5.4. Relationship of equilibrium constant (K_U) and free energy of unfolding (ΔG_U) with the spontaneity of the reaction		
K_U	ΔG_U	Direction
$K_U > 1$	$\Delta G_U < 0$	Favors unfolded state
$K_U < 1$	$\Delta G_U > 0$	Favors folded state
$K_U = 1$	$\Delta G_U = 0$	An equal mixture of folded and unfolded VH

At 30°C, ΔG_U and K_U for the VH-B1a domain was 9.60 ± 19.81 kcal/mol and 1.19 x 10^{-7}, and ΔG_U and K_U for the +VH domain was 11.00 ± 8.03 kcal/mol and 1.16 x 10^{-8}, showing that these VH domains favor the folded state at 30°C (Table 5.3 and 5.4).

5.1.6.3.3. Effect of framework mutations on VH domain stability

The change in +VH stability relative to VH-B1a (ΔΔG) describes the effect of framework mutations on VH domain stability. In 2008, Capriotti and Emidio, *et al.* implemented a predictor for the discrimination between 3 classes of mutations (destabilizing, neutral and stabilizing mutations) (Capriotti *et al.*, 2008). Each mutation can be categorized based on ΔΔG as described in Table 5.5.

Table 5.5. Mutation based on ΔΔG	
ΔΔG	Mutations
ΔΔG < -1.0 kcal/mol	Destabilizing mutation
ΔΔG > 1.0 kcal/mol	Stabilizing mutation
-1.0 ≤ ΔΔG ≤ 1.0 kcal/mol	Neutral mutation

We determined the extent of positively charged mutations on the stability of the VH-B1a by calculating the change in free energy (ΔΔG) for +VH domain relative to VH-B1a. The VH domain stability change of +VH domain: ΔΔG = ΔG(VH-B1a) - ΔG(+VH) kcal/mol = -1.4 ± 27.84 kcal/mol, which indicated that four framework mutations in VH-B1a domain have slightly destabilizing effect on the +VH domain stability (Table 5.3 and 5.5).

5.1.7. Solubility of VH-B1a and +VH domains after heating and cooling

We compared the amount of VH domain aggregation that occurred after heating VH domains to 90°C for 20 minutes, cooling them to 25°C, and centrifuging at 12,000 × g. The concentration of VH domain that remained in solution was calculated using the BCA protein

assay. Similar to the CD analysis of the fraction of VH domain folded after heating and cooling (Section 5.1.6.2), the fraction of soluble VH-B1a was higher relative to the +VH domain with a $p < 0.0001$, indicating the differences among means of VH-B1 and VH domains were statistically significant at $\alpha = 0.05$ (Figure 5.12).

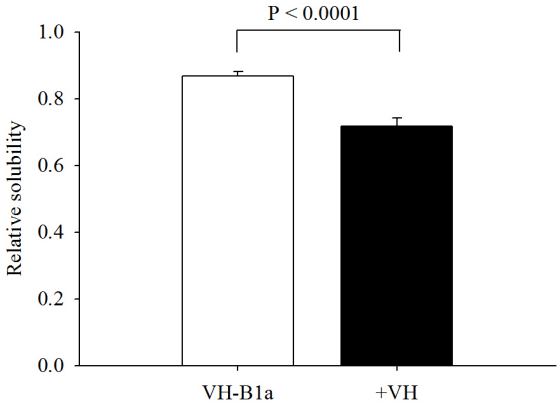


Figure 5.12. Solubility of VH-B1a and +VH domains after heating and cooling
Solubility analysis of VH-B1a and +VH domains after transient heating (90°C, 20 minutes), cooling (25°C) and sedimentation (12,000 × g). P-value was calculated using ordinary one-way ANOVA of Prism 7.0 (GraphPad).

5.1.8. Solubility of VH-B1a and +VH domains at room temperature

The long-term solubility of VH-B1a and +VH domains were measured by incubating VH domains at room temperature (25°C) for seven days. The concentration of soluble VH domain was assessed using the BCA protein assay after centrifuging the VH domain solutions at 10,000 × g. For long-term solubility at room temperature, the fraction of the soluble +VH domain was lower relative to VH-B1a with a $p < 0.0001$, indicating the differences among means of VH-B1 and VH domains were statistically significant at $\alpha = 0.05$ (Figure 5.13).

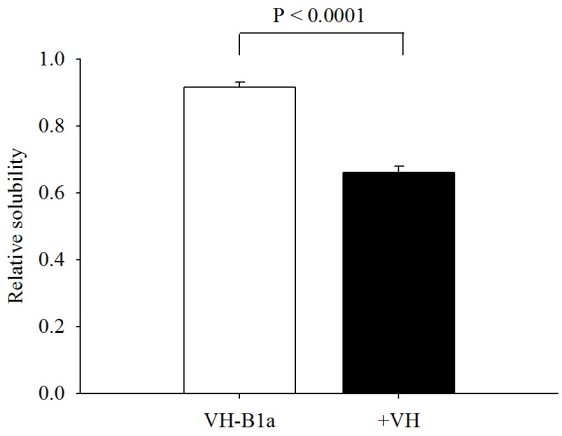


Figure 5.13. Solubility of VH-B1a and +VH domains at room temperature
Solubility analysis of VH-B1a and +VH domains at room temperature (25°C) after seven days. P-value was calculated using ordinary one-way ANOVA of Prism 7.0 (GraphPad).

5.2. Specific Aim 2: Assess the ability of VH domain framework net charge on tolerating charge in CDR1

5.2.1. Introduction

VH domains interact with antigen through their CDR loops, which is facilitated by the hydrophobic amino acid residues. Unfortunately, hydrophobic residues of the CDR loops promote VH domain aggregation. Several studies have explored the effect of CDR mutations in mediating VH domain aggregation (Jespers *et al.*, 2004; Arbabi-Ghahroudi *et al.*, 2009; Perchiacca *et al.*, 2011; Dudgeon *et al.*, 2012; Perchiacca *et al.*, 2012; Perchiacca *et al.*, 2014). Perchiacca and coworkers found that the CDR1 loop, but not the CDR2 and CDR3, can prevent VH domain aggregation. They designed three VH variants, each one displaying one CDR loop from the highly soluble antibody Hel4 and two CDR loops from the aggregation-sensitive Dp47d. They compared the solubility of the designed VH variants with the aggregation-prone Dp47d VH domain after heating to 95°C. They observed that the ability to resist aggregation was localized to the charged ^{31}DED^{33} residues of the CDR1 loop of Hel4, as only Dp47d bearing CDR1 from Hel4 folded reversibly. In contrast, CDR2 and CDR3 from Hel4 were unable to prevent aggregation (Perchiacca *et al.*, 2011). Moreover, Dudgeon and co-workers demonstrated that the insertion of DDD triple-charged residues in the 31-33 position of CDR1 loop provides superior biophysical properties to the VH domain, including increased protein expression, aggregation resistance, and reversible folding ability after heating. They noticed a detectable preference for aspartate over glutamate at all analyzed positions and a strong preference for aspartate over lysine or arginine (Dudgeon *et al.*, 2012). In another study, Perchiacca and co-workers observed that insertion of negatively charged residues near the edges of hydrophobic CDR is more efficient for inhibiting VH domain aggregation than positive charge for a VH scaffold that is negative or near-neutrally charged at neutral pH. They also found that insertion of positively charged residues near the edges of hydrophobic CDR is more efficient for inhibiting VH domain aggregation than negatively charged ones for a VH scaffold that is positively charged at neutral pH (Perchiacca *et al.*, 2014). Here, we investigated how the net framework charge of the VH domain facilitates CDR1 to tolerate charge residues for providing thermostability and solubility.

5.2.2. Construction of CDR1-mutated VH-B1a and +VH domains

To evaluate the effect of charge within CDR1 of VH-B1a and +VH domains, we designed ten CDR1-mutated VH domains where the ^{31}DTY^{33} sequence of CDR1 was replaced with a series of single, double, and triple positively and negatively charged amino acid residues. For CDR1-mutated VH-B1a domains, their theoretical pIs cover a range from 6.79 to 9.06 and theoretical net charge from -1.70 to 4.30 at pH 7.4. CDR1-mutated +VH domains have theoretical pIs ranging from 8.45 to 9.60 and theoretical net charges from 2.30 to 8.30 at pH 7.4 (http://protcalc.sourceforge.net/). For all CDR1-mutated VH domains, CDR2 and CDR3 were not changed from VH-B1a (Table 5.6).

Table 5.6. Designed CDR1-mutated VH-B1a and +VH domains

CDR1-mutated VH domains	CDR1 mutation (^{31}DTY^{33})	Net CDR1 charge	pI and net charge of CDR1-mutated VH-B1a domains at pH 7.4	pI and net charge of CDR1-mutated +VH domains at pH 7.4
DDD	31-33DDD	-3.20	6.79, -1.70	8.45, 2.30
DDY	31, 32DD	-2.20	7.09, -0.70	8.76, 3.30
DTD	31, 33DD	-2.20	7.09, -0.70	8.78, 3.30
DTY	31D	-1.20	7.56, 0.30	9.04, 4.30
DKY	32K	-0.20	8.03, 1.30	9.20, 5.30
DTK	33K	-0.20	8.03, 1.30	9.24, 5.30
DKK	32, 33KK	0.80	8.45, 2.30	9.38, 6.30
KTY	31K	0.80	8.45, 2.30	9.38, 6.30
KKY	31,32KK	1.80	8.78, 3.30	9.48, 7.30
KTK	31,33KK	1.80	8.80, 3.30	9.51, 7.30
KKK	31-33KKK	2.80	9.06, 4.30	9.60, 8.30

To construct VH domain CDR1 mutants, we made dU-ssDNA from VH-B1a and +VH domain cloned in the pET22b(+) vector (Section 4.5.1). We used the Kunkel mutagenesis method (Kunkel *et al.*, 1987) with the dU-ssDNA plasmid and mutagenic oligonucleotides 10-19 (Table 4.5) to replace ^{31}DTY^{33} sequence from the CDR1 loop with charged residues (Section 4.5.2). Purified Kunkel mutation reaction products were used to transform in *E. coli* DH10B (Section 4.3.5). Plasmids were amplified and purified as per Section 4.3.3. Following sequence verification (Section 4.2.4), plasmids were used for VH domain protein expression (Section 4.4.2 and 4.4.3) and purification (Section 4.4.4).

5.2.3. Expression optimization of CDR1-mutated VH-B1a and +VH domains

We tried to express CDR1-mutated VH-B1a and +VH domains in the BL21 bacterial strain using autoinduction media at 30°C for 24 hours. VH-B1a domains with negatively charged mutations and a single lysine mutation in the CDR1 expressed in the culture supernatant. We were unable to express VH-B1a domains with double and triple lysine residues in the CDR1 and all CDR1-mutated +VH domains under these conditions.

We were able to express these VH domains in the culture supernatant when they were expressed in 2YT media with 1 mM IPTG induction for 48 hours at 16°C (Section 4.4.3). The expression level of VH-B1a domains was increased with increasing numbers of negatively charged residues in the CDR1 loop, ranging from 20 mg/L to 27 mg/L and decreased with increased number of lysine mutations ranging from 20 mg/L to 8 mg/L. Similarly, expression levels of +VH domains increased with increasing number of negatively charged residues ranging from 15 mg/L to 20 mg/L and decreased with increasing number of lysine mutations ranging from 15 mg/L to 10 mg/L (Table 5.7).

Table 5.7. Effect of CDR1 mutations on protein expression of CDR1-mutated VH-B1a and +VH domains

CDR1-mutated VH domains	VH-B1a domains Yield (mg/l)[a]	+VH domains Yield (mg/l)[a]
DDD	27	20
DDY	23	17
DTD	22	16
DTY	20	15
DKY	13	13
DTK	13	12
DKK	11	14
KTY	12	14
KKY	11	12
KTK	10	12
KKK	8	10

[a]Yield (mg/l) was calculated using 200 ml bacterial cultures

5.2.4. Purification of CDR1-mutated VH-B1a and +VH domains

After optimization, all VH domains were expressed in 2YT media using the IPTG-based induction method (Section 4.4.3), purified with His-Tag binding resin (section 4.4.4), and analyzed by SDS-PAGE (Section 4.4.5 and 4.4.6) (Figure 5.14).

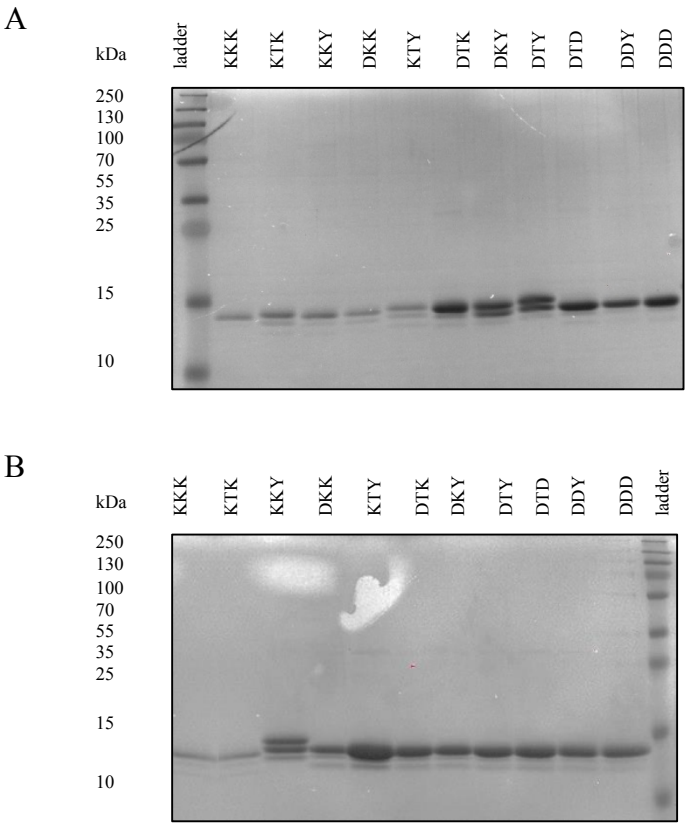


Figure 5.14. SDS-PAGE of purified CDR1-mutated VH-B1a and +VH domains
CDR1-mutated VH domains were expressed in the BL21 strain using 2YT media with 1 mM IPTG induction and purified by His-based purification method. Purified VH-B1a (A) and +VH (B) domains were run on an SDS-PAGE gel and visualized by Coomassie staining. ^{31}DTY^{33} is the wild type CDR1 sequence of VH-B1a and +VH domain, and the other CDR1-mutated VH-B1a and +VH domains having a series of aspartate (D) and lysine (K) mutations at 31-33 position of CDR1 loop.

5.2.5. Secondary structure analysis of the CDR1-mutated VH-B1a and +VH domains

We evaluated secondary structural patterns of CDR1-mutated VH-B1a and +VH domains using circular dichroism spectroscopy. Circular dichroism analysis revealed that CDR1-charged mutants possess similar secondary structure to VH-B1a and +VH domain, indicating that CDR1 mutations did not substantially change the structure of VH-B1a and +VH domains (Figure 5.15).

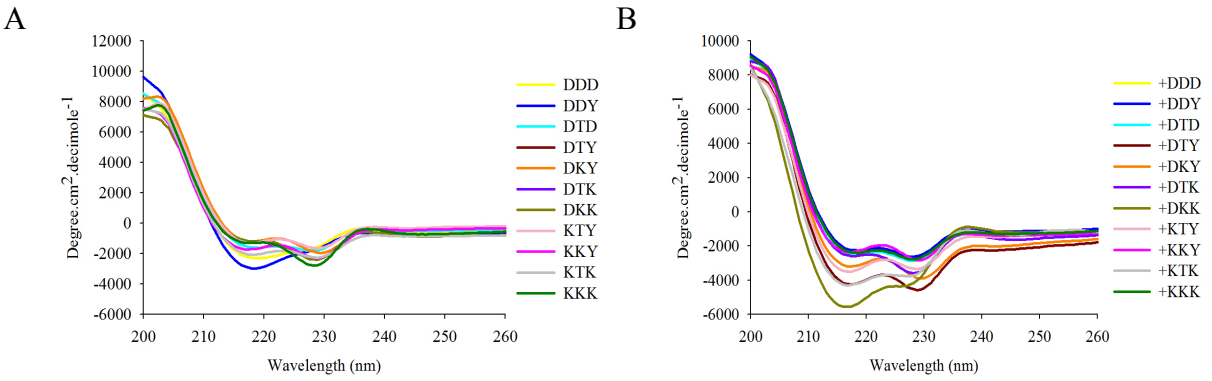


Figure 5.15. Secondary structure of the CDR1-mutated VH-B1a and +VH domains
Far-UV CD spectra of VH-B1a (A) and +VH (B) domains at 30°C. Color lines represent CD spectra of CDR1-mutated VH-B1a and +VH domains at 30°C. CDR1 mutations are listed to the right.

5.2.6. Thermostability of the CDR1-mutated VH-B1a and +VH domains

We measured the thermostability of CDR1-mutated VH-B1a and +VH domains, as described in Section 5.1.6.2 and calculated their melting temperature (Tm) from temperature-induced denaturation curves (Figure 5.16). We measured Tms and the fraction of correctly folded VH domains after heating to 90°C and cooling. After cooling we reheated the VH domains to calculate the Tm and the fraction of correctly folded VH domain for a second time (Figure 5.16).

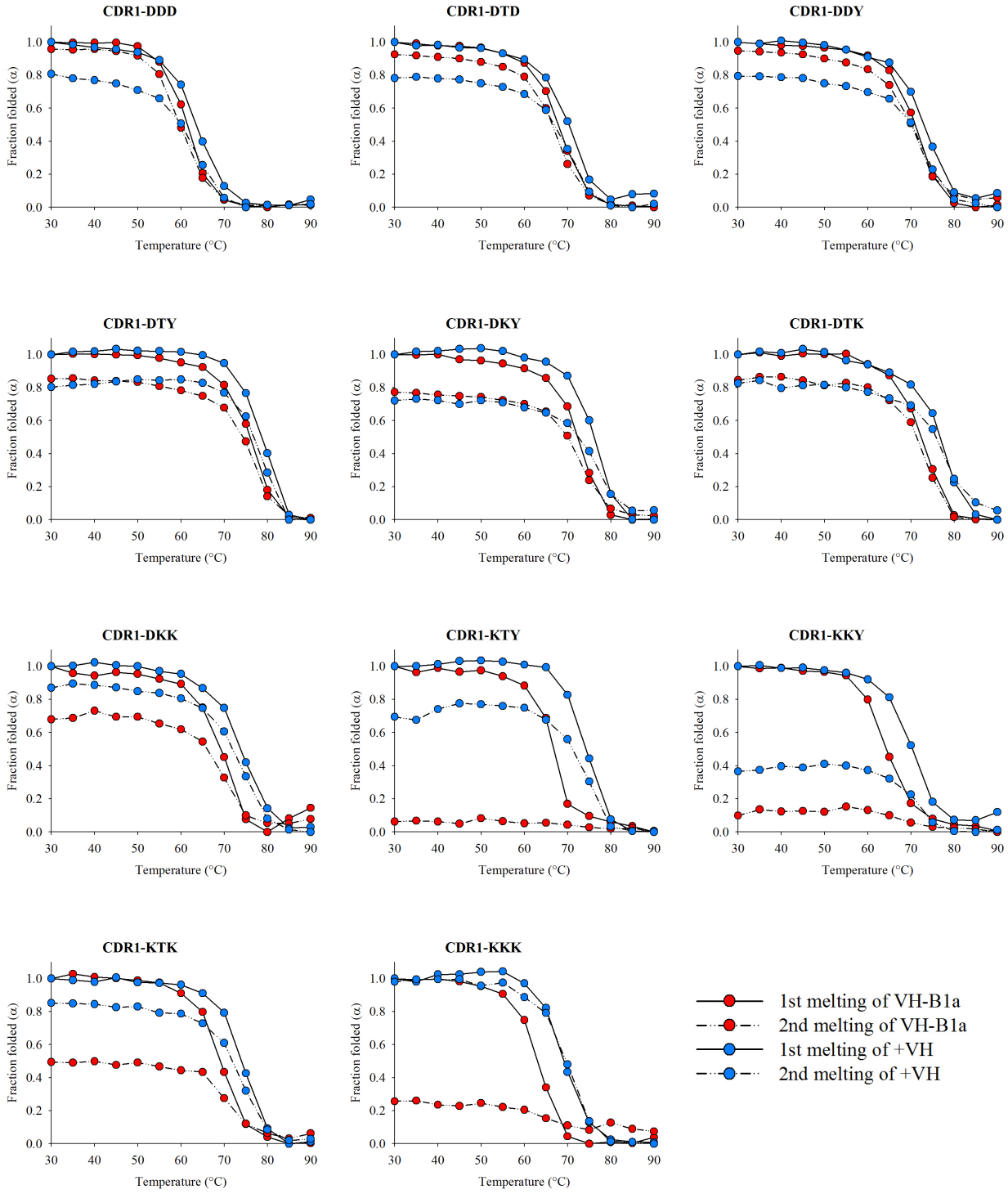


Figure 5.16. Temperature induced denaturation of CDR1-mutated VH-B1a and +VH domains

CD analysis of the first (solid line) and second (dashed line) melting of CDR1-mutated VH-B1a and +VH domains at 207 nm. Red circle: first and second melting of CDR1-mutated VH-B1a domains; blue circle: first and second melting of CDR1-mutated +VH domains.

5.2.6.1. Effect of CDR1 net charge on thermostability of the CDR1-mutated VH-B1a and +VH domains

We examined how the melting temperature (Tm) of CDR1-mutated VH-B1a and +VH domains changed in response to the CDR1 net charge (Table 5.8).

Table 5.8. Effect of CDR1 mutation on thermostability of CDR1-mutated VH-B1a and +VH domains at pH 7.4

CDR1-mutated VH domains	VH-B1a domains Tm (°C) ± SEE[a]	+VH domains Tm (°C) ± SEE[a]	ΔTm (°C) (Tm of +VH – Tm of VH-B1a)
DDD	61.50 ± 0.01	63.50 ± 0.02	2.00
DDY	71.00 ± 0.02	73.00 ± 0.03	2.00
DTD	68.00 ± 0.02	70.00 ± 0.04	2.00
DTY	76.00 ± 0.03	78.00 ± 0.03	2.00
DKY	72.50 ± 0.03	76.00 ± 0.03	3.50
DTK	73.00 ± 0.02	77.00 ± 0.04	4.00
DKK	70.00 ± 0.06	74.00 ± 0.02	4.00
KTY	67.00 ± 0.04	73.50 ± 0.02	6.50
KKY	64.50 ± 0.02	70.00 ± 0.04	5.50
KTK	69.00 ± 0.02	74.00 ± 0.02	5.00
KKK	63.00 ± 0.02	69.00 ± 0.02	6.00

[a]Standard Error of Estimate (SEE) was calculated using non-linear regression analysis with SigmaPlot 10.0

We plotted the net CDR1 charge versus Tm for VH-B1a and +VH domains to compare how CDR1 net charge influenced Tm (Figure 5.17). Tms decreased as negative and positive charge was added to the DTY CDR1 sequence present in VH-B1a and +VH, producing a bell-shaped curve. In this curve, we observed a trend whereby a sudden drop of Tm occurred for the VH-B1a domain due to a position specific effect of a single lysine in the 31-position and a double lysine in (31, 32)-positions of CDR1. The decrease in Tm due to the single lysine in 31-position was less for the +VH domain. However, the decrease in Tm caused by the double lysine in (31, 32)-positions was still evident in the +VH domain.

These results showed that CDR1 charge has a large effect on Tm for both VH-B1a (61.50 to 76°C) and +VH (63.50 to 78°C) domains. The +VH domain had higher Tm values than VH-B1a for both positive and negative CDR1 mutations. The difference in Tm values between the VH-B1a and +VH domains were minimal for mutations with negative CDR1 charge (range of 2°C) and a larger difference was observed for positive charge mutations (range

of 3.50 to 6.50°C), where the +VH domain was better able to accommodate positive charge in CDR1 (Figure 5.17B).

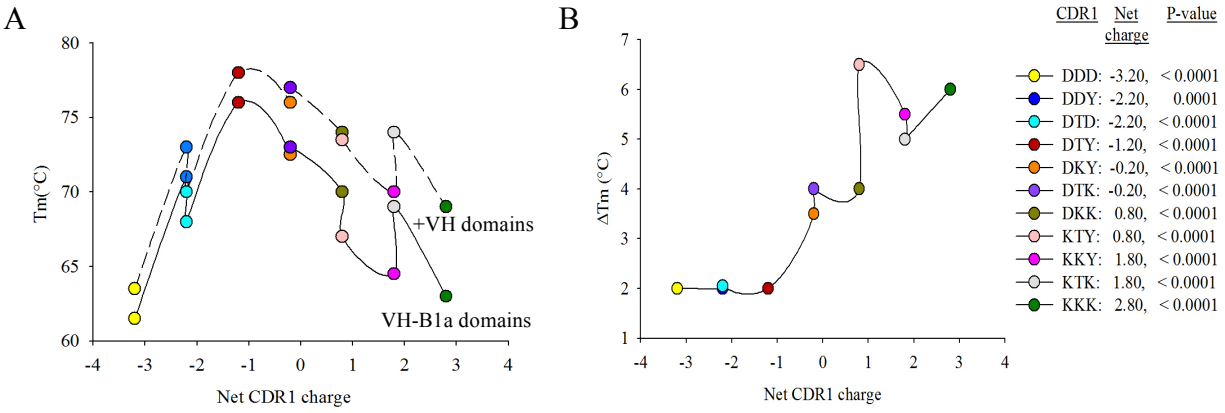


Figure 5.17. Effect of net CDR1 charge on thermostability of CDR1-mutated VH-B1a and +VH domains

(A) Melting temperature (Tm) of CDR1-mutated VH-B1a (solid line), and +VH domains (dashed line) were calculated from the CD data monitored at 207 nm during first melting. (B) ΔTm is the Tm difference between +VH and VH-B1a domain. Color dots represent CDR1-mutated VH-B1a (solid line) and +VH domains (dashed line). CDR1 mutations with their net charge, and p-values between VH-B1a and +VH domains for specific CDR1 are listed to the right. P-values were calculated from curve fitting with a Boltzmann sigmoidal curve and compared with Prism 7.0 (GraphPad).

5.2.7. Fraction of VH-B1a and +VH domains folded with mutated CDR1

We measured the ability of the CDR1-mutated VH-B1a and +VH domains to reversibly fold by measuring the fraction of VH domain folded after heating to 90°C and cooling to 30°C (Figure 5.16) (Table 5.9).

Table 5.9. Effect of CDR1 mutations on fraction of CDR1-mutated VH-B1a and +VH domains folded at pH 7.4

CDR1-mutated VH domains	VH-B1a domains Fraction folded (α) ± SEE[a]	+VH domains Fraction folded (α) ± SEE[a]
DDD	0.96 ± 0.01	0.77 ± 0.02
DDY	0.95 ± 0.03	0.80 ± 0.03
DTD	0.93 ± 0.02	0.78 ± 0.02
DTY	0.85 ± 0.03	0.77 ± 0.02
DKY	0.77 ± 0.02	0.72 ± 0.02
DTK	0.85 ± 0.03	0.83 ± 0.02
DKK	0.68 ± 0.03	0.87 ± 0.02
KTY	0.06 ± 0.01	0.70 ± 0.03
KKY	0.10 ± 0.02	0.36 ± 0.02
KTK	0.49 ± 0.02	0.85 ± 0.02
KKK	0.26 ± 0.02	0.98 ± 0.02

[a]Standard Error of Estimate (SEE) was calculated using non-linear regression analysis with SigmaPlot 10.0

We compared the fraction of folded (α) CDR1-mutated VH-B1a and +VH domains versus net CDR1 charge to observe how CDR1 charge affected the reversible folding ability of VH-B1a and +VH domains. The net CDR1 charge significantly affected the reversible folding ability of both VH-B1a (6% to 96%) and +VH (36% to 98%) domains. For CDR1-mutated VH-B1a domains, we observed a decrease in the reversible folding as the positive charge in CDR1 was increased. For +VH domains, we found a slight increase in the reversible folding as the positive charge in CDR1 increased. In general, VH-B1a had the higher reversible folding ability for negative CDR1 mutations. Conversely, +VH domain showed higher reversible folding ability for positive CDR1 mutations. This preference indicated that negatively charged aspartate in CDR1 prevented VH-B1a aggregation and positively charged lysine promoted aggregation. Moreover, +VH better tolerated positive and negative mutations in the CDR1. Overall, the difference in reversible folding between the VH-B1a and +VH domains were minimal for negative CDR1 mutations and larger for positive CDR1 mutations (Figure 5.18).

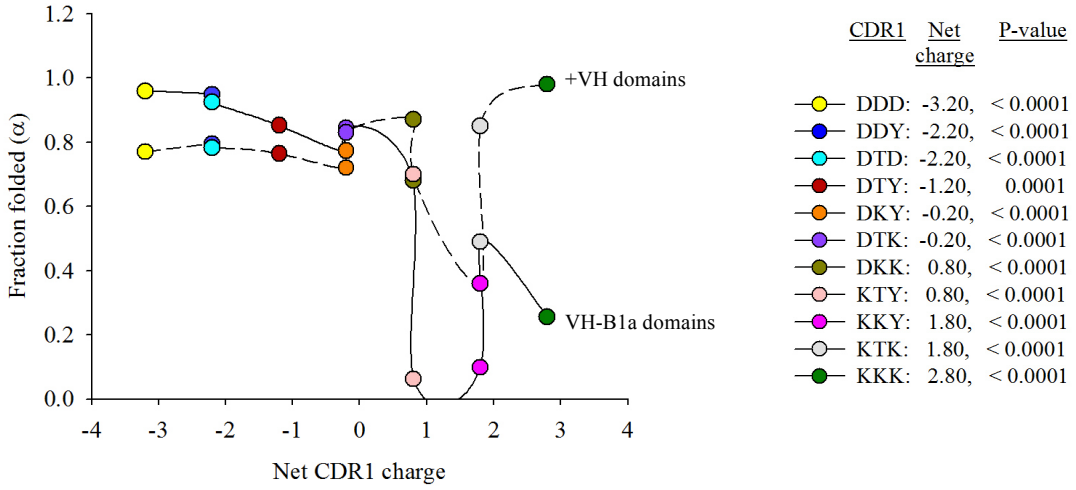


Figure 5.18. Effect of net CDR1 charge on reversible folding ability of CDR1-mutated VH-B1a and +VH domains
Fraction folded (α) of CDR1-mutated VH-B1a (solid line), and +VH domains (dashed line) were calculated from the CD data monitored at 207 nm during first and second melting. Color dots represent CDR1-mutated VH-B1a (solid line) and +VH domains (dashed line). CDR1 mutations with their net charge, and p-values between VH-B1a and +VH domains for specific CDR1 are listed to the right. P-values were calculated from curve fitting with a Boltzmann sigmoidal curve and compared with Prism 7.0 (GraphPad).

We observed some fluctuations in the trend of CDR1 charge on the fraction of folded VH-B1a and +VH domains (Figure 5.18). We observed position specific effects of CDR1 mutations that reduced the ability of the VH domains to refold after heating. For example, a single lysine in 31-position and double lysine in (31, 32)-positions of the CDR1 reduced the reversible folding ability of VH-B1a. The position specific effect of single lysine in 31-position on the amount of folded VH-B1a domain was decreased in the +VH domain. We observed the positional effect of double lysine in (31, 32)-positions of the CDR1 was better tolerated in +VH domain compared to VH-B1a.

5.2.8. Temperature-induced aggregation of CDR1-mutated VH-B1a and +VH domains

To determine how resistant the CDR1-mutated VH domains were to aggregation after heating, we heated the CDR1-mutated VH domains from 30-90°C for 20 minutes and then cooled them to 25°C, and measured the concentration of soluble protein (Figure 5.19).

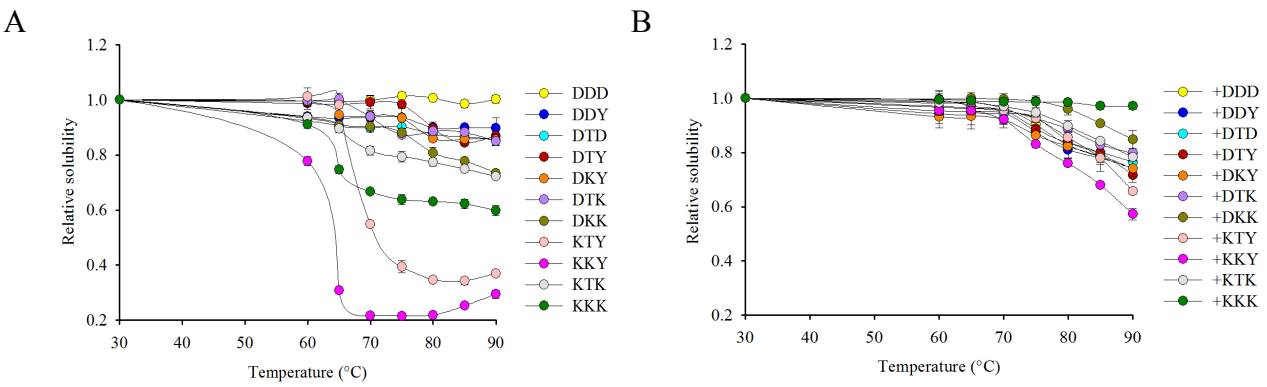


Figure 5.19. Effect of CDR1 mutation on temperature induced aggregation of VH-B1a and +VH domains

Solubility analysis of the CDR1-mutated VH-B1a (A) and +VH (B) domains after transient heating (30-90°C, 20 minutes), cooling (25°C), and sedimentation (12,000 × g). Color dots represent CDR1-mutated VH-B1a and +VH domains. CDR1 mutations are listed to the right.

We summarized the relative solubility of the VH domains by comparing the solubility of the VH-B1a and +VH domains at 90°C with their solubility at 30°C (Table 5.10).

Table 5.10. Effect of CDR1 mutations on temperature-induced aggregation of CDR1-mutated VH-B1a and +VH domains at pH 7.4

CDR1-mutated VH domains	VH-B1a domains Relative solubility at 90°C ± Standard Error	+VH domains Relative solubility at 90°C ± Standard Error
DDD	1.00 ± 0.01	0.74 ± 0.01
DDY	0.90 ± 0.04	0.74 ± 0.02
DTD	0.86 ± 0.02	0.77 ± 0.02
DTY	0.87 ± 0.01	0.72 ± 0.03
DKY	0.86 ± 0.02	0.74 ± 0.03
DTK	0.85 ± 0.01	0.80 ± 0.01
DKK	0.73 ± 0.01	0.85 ± 0.03
KTY	0.37 ± 0.01	0.66 ± 0.01
KKY	0.29 ± 0.02	0.57 ± 0.02
KTK	0.72 ± 0.01	0.78 ± 0.02
KKK	0.60 ± 0.02	0.97 ± 0.01

The VH-B1a domain showed a slight decrease in solubility after heating as the CDR1 increased from net negative to positive charge. We observed position specific effects of CDR1 mutations that reduced the solubility of VH domains after heating. For example, a single lysine in 31-position and double lysine in (31, 32)-positions of the CDR1 reduced the solubility of

VH-B1a. For +VH domain, we observed a slight increase in solubility as the net charge in CDR1 changed from negative to positive. Similar to the VH-B1a domain, we observed the same CDR positional effects that reduced solubility.

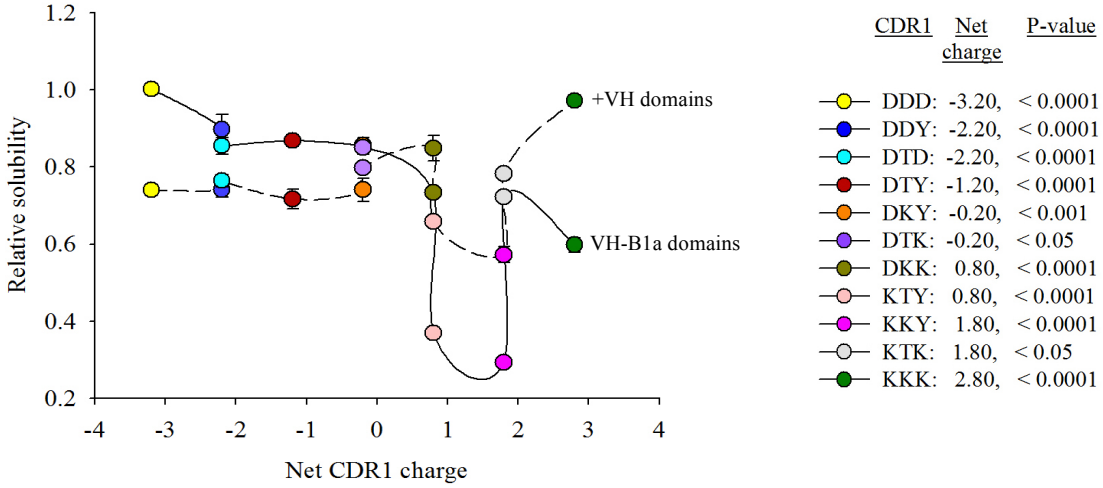


Figure 5.20. Effect of net CDR1 charge on temperature induced aggregation of CDR1-mutated VH-B1a and +VH domains
Solubility analysis of the CDR1-mutated VH-B1a and +VH domains after transient heating (90°C, 20 minutes), cooling (25°C) and sedimentation (12,000 × g). Color dots represent CDR1-mutated VH-B1a (solid line) and +VH domains (dashed line). CDR1 mutations with their net charge, and p-values between VH-B1a and +VH domains for specific CDR1 are listed to the right. P-values were calculated from multiple comparisons using ordinary one-way ANOVA of Prism 7.0 (GraphPad).

These trends in solubility changes after heating were similar to the change in fraction folded after heating measured using CD (Figure 5.18). We plotted the fraction of VH domain folded after heating versus the relative solubility after heating to determine the correlation between these two parameters (Figure 5.21). We observed a positive correlation between the amount of folded VH domain and the solubility after heating. VH domains with low ability to refold after heating also had low solubility. As the fraction re-folded increased so did solubility.

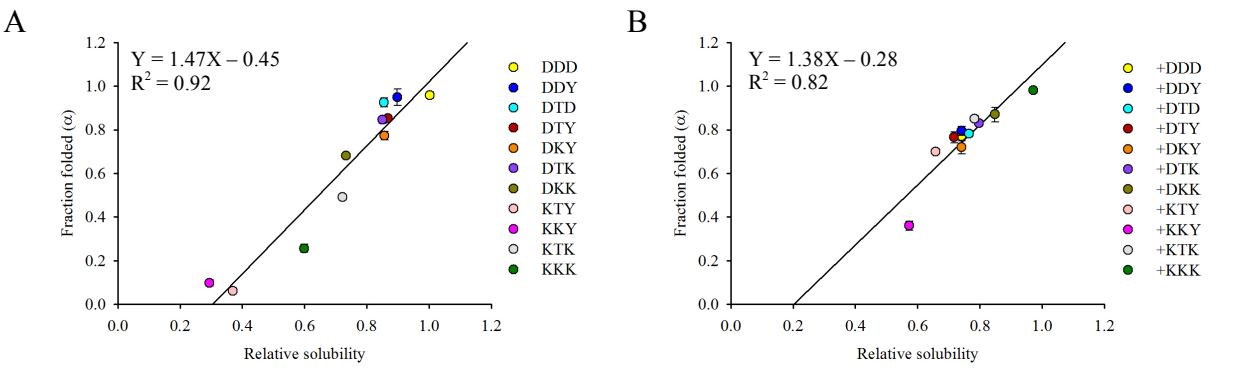


Figure 5.21. Fraction of CDR1-mutated VH-B1a and +VH domains folded are positively correlated with their solubility after heating
Linear regression analysis of fraction folded and solubility after heating of CDR1-mutated VH-B1a (A) and +VH (B) domains with SigmaPlot. R^2 represents the strength of relationship between fraction folded and solubility at 90°C. Color dots represent CDR1-mutated VH-B1a and +VH domains. CDR1 mutations are listed to the right.

5.2.9. Solubility of CDR1-mutated VH-B1a and +VH domains at room temperature

The long-term room temperature solubilities of CDR1-mutated VH-B1a, and +VH domains were measured by incubating them at room temperature (25°C) for seven days. VH domain samples were collected after every 24 hours and aggregates were removed by centrifuging at 10,000 × g. Soluble VH domain concentration was calculated using the BCA protein assay (Figure 5.22) (Table 5.12).

The room temperature solubility of CDR1-mutated VH-B1a domains decreased compared to VH-B1a with DTY CDR1, except for the DTK CDR1, which showed the highest solubility from day 1 to day 7. On the other hand, KKY, KTK, and DDD were much less soluble, and KKK started to aggregate after three days (Figure 5.22A). The room temperature solubility of CDR1-mutated +VH domains decreased compared to +VH with DTY CDR1. However, +DDD and +KKK started aggregating after four days (Figure 5.22B).

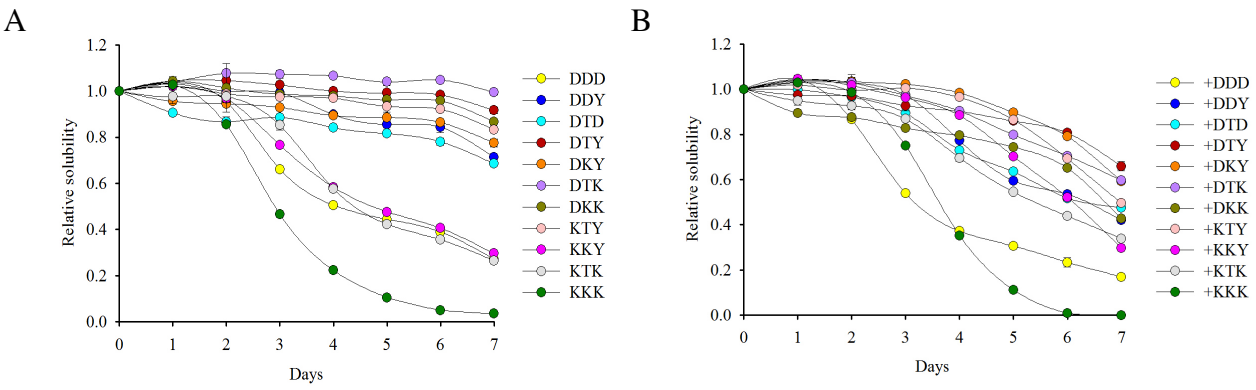


Figure 5.22. Effect of CDR1 mutation on room temperature solubility of VH-B1a and +VH domains

Solubility analysis of the CDR1-mutated VH-B1a (A) and +VH (B) domains at room temperature (25°C) for seven days. Color dots represent CDR1-mutated VH-B1a and +VH domains. CDR1 mutations are listed to the right.

We summarized the relative solubility of the VH domains by comparing the solubility of the VH-B1a and +VH domains at day seven with their solubility at day zero (Table 5.11).

Table 5.11. Effect of CDR1 mutations on room temperature solubility of CDR1-mutated VH-B1a and +VH domains at pH 7.4

CDR1-mutated VH domains	VH-B1a domains Relative solubility at day seven	+VH domains Relative solubility at day seven
DDD	0.27 ± 0.01	0.17 ± 0.01
DDY	0.71 ± 0.01	0.42 ± 0.01
DTD	0.69 ± 0.01	0.48 ± 0.01
DTY	0.92 ± 0.01	0.66 ± 0.02
DKY	0.78 ± 0.02	0.59 ± 0.01
DTK	0.99 ± 0.01	0.60 ± 0.01
DKK	0.87 ± 0.01	0.43 ± 0.01
KTY	0.83 ± 0.01	0.50 ± 0.01
KKY	0.30 ± 0.01	0.30 ± 0.01
KTK	0.26 ± 0.01	0.34 ± 0.01
KKK	0.03 ± 0.01	0.00 ± 0.02

We plotted room temperature solubility of CDR1-mutated VH-B1a and +VH domains with net CDR1 charge to examine how the solubility of CDR1-mutated VH-B1a and +VH domains changed with the net CDR1 charge. We did not observe a linear trend of solubility with net CDR1 charge, rather we obtained a bell-shaped curve, where the room temperature

solubility of VH-B1a and +VH domains decreased with increasing or decreasing net CDR1 charge. CDR1 charge had a large effect on the room temperature solubility of VH-B1a (a range of zero to 99%) and +VH domain (zero to 66%). Overall, the VH-B1a domain had higher solubility than +VH for both positive and negative CDR1 mutations (Figure 5.23).

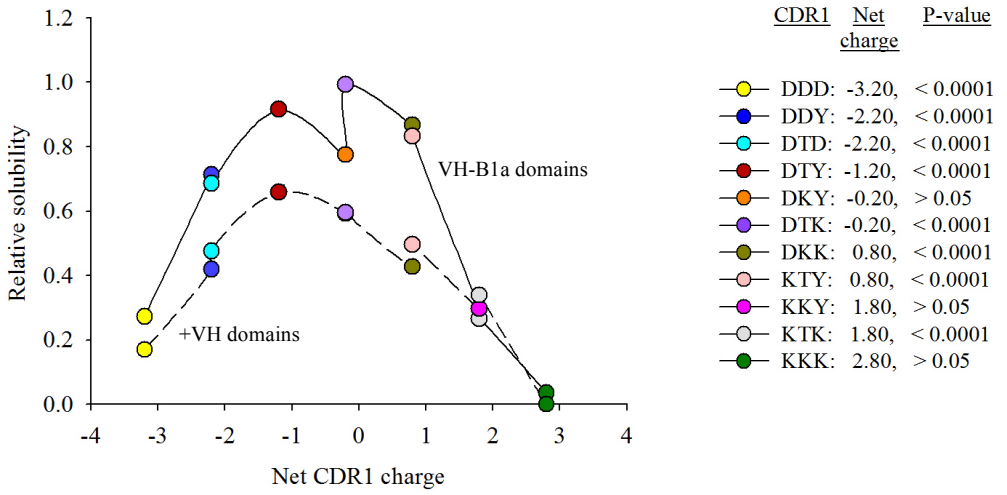


Figure 5.23. Effect of net CDR1 charge on room temperature solubility of CDR1-mutated VH-B1a and +VH domains
Solubility analysis of the CDR1-mutated VH-B1a and +VH domains at room temperature (25°C) after seven days. Color dots represent CDR1-mutated VH-B1a (solid line) and +VH domains (dashed line). CDR1 mutations with their net charge, and p-values between VH-B1a and +VH domains for specific CDR1 are listed to the right. P-values were calculated from multiple comparisons using ordinary one-way ANOVA of Prism 7.0 (GraphPad).

The trend in room temperature solubility verses net CDR1 charge was similar to the trend in Tm versus net CDR1 charge (Figure 5.17). We plotted room temperature solubility versus Tm for VH domains to determine the correlation between these two parameters (Figure 5.24). From the linear regression analysis of the plotted Tm (°C) versus room temperature solubility, we observed that the room temperature solubility of CDR1-mutated VH-B1a and +VH domains were positively correlated with the melting point (Tm). The results show that the higher the Tm of a VH domain the higher its room temperature solubility.

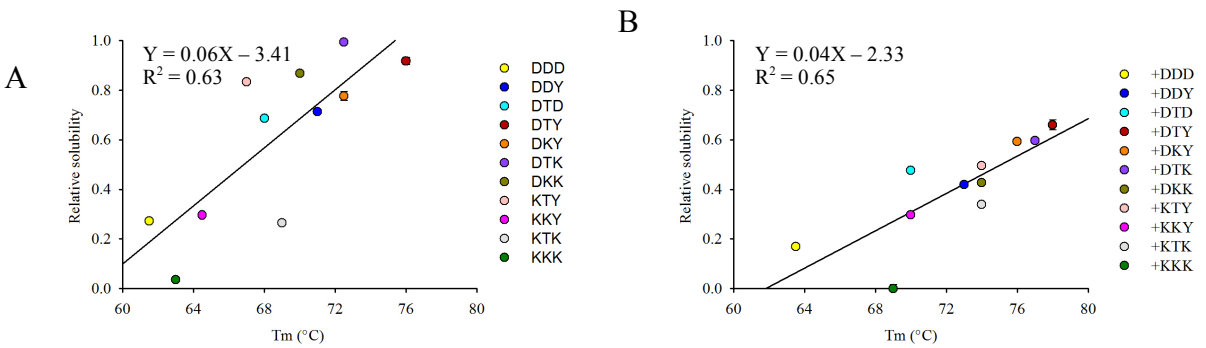


Figure 5.24. Tm of CDR1-mutated VH-B1a and +VH domains are positively correlated with their room temperature solubility

Linear regression analysis of Tm and room temperature solubility of CDR1-mutated VH-B1a (A) and +VH (B) domains with SigmaPlot. R^2 represents the strength of the relationship between Tm and room temperature solubility. Color dots represent CDR1-mutated VH-B1a and +VH domains. CDR1 mutations are listed to the right.

6. DISCUSSIONS AND FUTURE DIRECTIONS

Antibodies are widely used in diagnostic and therapeutic applications. The large size and tetrameric structure of antibodies limit their penetration through the cell membrane for intracellular applications as well as inhibits tissue penetration, limiting its usefulness in therapeutic applications. Moreover, antibodies cannot fold correctly in the reducing environment of bacterial cytoplasm, which decreases their production in bacterial systems. For diagnostic and therapeutic purposes, there is a need to produce antibodies with high quality and in sufficient quantity, however, production needs to be cost-effective. Since the full antibody is not required for target binding, smaller antibody fragments called nanobodies (VH and VL) have been developed for intracellular and therapeutic applications. Nanobodies retain the antigen binding affinity of antibodies, but lack the effector function of the Fc region. The variable heavy (VH) domain is the most commonly used nanobody as VH is more thermostable than variable light (VL) domain (Rothlisberger *et al.*, 2005). Moreover, VH domains are observed in nature, such as in camelids (Hamers-Casterman *et al.*, 1993) and cartilaginous fish (Hinds and Litman, 1986). The development of VH domain circumvents many of the issues associated with antibodies. However, they are not without their unique set of limitations and obstacles. In the absence of an interacting VL domain, the hydrophobic residues of VH/VL interface are exposed to the external environment that facilitates VH domain aggregation. Despite this intrinsic instability of human VH domains, several research groups were able to engineer VH domains with enhanced stability and solubility. In this thesis, we explored whether VH stability and solubility can be improved through charge engineering of VH framework and complementarity determining regions (CDRs).

Although CDRs are key determinants of VH aggregation, the VH framework is also a significant contributor to their solubility. VH solubility can be increased by altering the net charge of the VH scaffold (Perchiacca *et al.*, 2011, 2012; Miklos *et al.*, 2012). Lawrence *et al.* showed that proteins are responsive to radical changes of their surface charge, which leads to enhanced solubility and reduced aggregation (Lawrence *et al.*, 2007). Further, Miklos's group used "supercharging" strategy to enhance scFv solubility, where they mutated the solvent-exposed residues of the framework region to the charge residues of the same polarity (Miklos *et al*., 2012).

In this thesis, we assessed the influence of framework net charge on thermostability, reversible folding, and solubility of VH domain. We chose a thermodynamically stable and near neutrally charged (+0.3 at pH 7.4) human VH-B1a domain as a starting point designed by Barthelemy's group (Barthelemy *et al.*, 2008). They replaced exposed hydrophobic residues of VH-4D5 with five structurally compatible mutations (H35G, Q39R, L45E, R50S, and S93A). From there, we followed the anti-MS2 scFv designed by Miklos *et. al.* (Miklos *et al*., 2012) to develop a positively (+4.30 at pH 7.4) and a negatively (-5.70 at pH 7.4) charged VH domain. We introduced four positive (S21K, A23K, A40R and T68K) and six negative (V5D, L11E, Q13D, S21E, A23D, and S70D) charged residues in the framework region of the VH-B1a for the construction of +VH and -VH domain, respectively. We observed a lower level of expression for +VH (16 mg/L) and -VH domains (3 mg/L) compared to VH-B1a (20 mg/L). Initially, we observed the secondary structural pattern of VH-B1a, +VH, and -VH from Circular Dichroism (CD) spectra. Four framework mutations of the +VH domain did not perturb the secondary structure, whereas the six framework mutations of -VH domain likely denatured the VH domain, therefore we excluded -VH from further experiments. The increased net positive charge of VH framework increased the thermostability from 76°C to 78°C, but lowered the reversible folding ability after denaturation, heat based solubility at 90°C and long-term solubility at room temperature. We also measured the thermodynamic parameters (ΔH, ΔS, and ΔG) of VH-B1a and +VH domains from CD data following the thermal unfolding curve, Van't Hoff plot and direct fit approach, and observed that four positive framework mutations have no thermodynamic effect on VH domain stability.

Another strategy that we employed to increase the stability and solubility of VH domain was to introduce charge mutations in the complementarity determining region (CDR) of VH domain. VH domains having identical framework regions, but different CDRs, show significant difference in VH stability and solubility (Perchiacca *et al.*, 2011). This indicates that CDRs are directly involved in mediating VH aggregation. Perchiacca and coworkers observed that the CDR1 loop, but not the CDR2 and CDR3, can prevent VH aggregation (Perchiacca *et al.*, 2011). Moreover, Dudgeon and co-workers demonstrated that the insertion of DDD triple-charged residues in the 31-33 position of CDR1 provides superior biophysical properties to the VH domain (Dudgeon *et al.*, 2012). In another study, Perchiacca and coworkers observed that the net charge of VH scaffold is the key determinant of optimal CDR mutation for preventing

VH aggregation (Perchiacca *et al.*, 2014). Together, these studies highlight the importance of optimal protein charge and proper charge distribution in preventing aggregation. From these studies, we introduced a series of positive and negative charge mutations in the CDR1 of VH-B1a and +VH domains to assess how the VH domain framework net charge can tolerate CDR1 charge and provide thermostability, reversible folding ability, and solubility. We designed ten CDR1-mutated VH-B1a (pI range: 6.79 to 9.06 and net charge: -1.70 to 4.30 at pH 7.4) and +VH domains (pI range: 8.45 to 9.60 and net charge: 2.30 to 8.30 at pH 7.4) to assess the ability of the VH domain framework net charge on tolerating charge in CDR1.

The expression of CDR1-mutated VH-B1a and +VH domains increased with increasing number of negatively charged aspartate residues and decreased with increasing number of lysine residues in the CDR1. However, CDR1 mutations did not have any effect on the secondary structure of VH-B1a and +VH domain.

The net charge of CDR1 has a significant impact on thermostability, reversible folding ability, temperature-induced aggregation and long-term solubility at room temperature. We determined the melting temperature (Tm) of the VH-B1a and +VH domain CDR1 charge mutants. The Tms ranged from 61.50 to 76°C and 63.50 to 78°C for the CDR1-mutated VH-B1a and +VH domains. For both VH domains, the DTY CDR1 had the highest Tm and adding negative or positive charge to the CDR1 decreased the Tm. We observed a positional effect of lysine mutations on the Tm. The VH-B1a domain with single lysine in 31-position and double lysine in (31, 32)-positions of CDR1 caused a dramatic reduction in Tm. The +VH domain was able to suppress the decreased Tm caused by the single lysine in 31-position, but not the decreased Tm caused by the double lysine in (31, 32)-positions in CDR1. Moreover, +VH always showed higher thermostability for positive and negative CDR1 mutations compared to VH-B1a. The thermostability difference between VH-B1a and +VH domains were minimal for negative charge mutations and higher for positive mutations in the CDR1; indicating that +VH has better tolerance for both negative and positive CDR1 charge, whereas VH-B1a can better accommodate negative charge.

We measured the ability of VH domains to become refolded after heat-induced denaturation. CDR1-mutated VH-B1a domains showed a decreased trend, and +VH domains showed an increased trend of reversible folding ability with increased net CDR1 charge. Similar to thermostability, we noticed position specific effect for single lysine in 31-position

and double lysine in (31, 32)-positions of CDR1 for both VH-B1a and +VH domains. However, this position specific effect could be overcome mostly when these mutations were introduced into the +VH domain, indicating that +VH can better tolerate the positively charged lysine in CDR1. Moreover, designed +VH can better tolerate both positive and negative charge in CDR1, whereas VH-B1a can better accommodate negative charge in CDR1 for high reversible folding ability. Hence, the difference in reversible folding ability was lower between VH-B1a and +VH domains for negative CDR1 mutations and significantly higher for positive CDR1 mutations.

We determined temperature-induced aggregation to determine how much VH domains become aggregated after denaturation. We observed a trend of temperature-induced aggregation of CDR1-mutated VH-B1a and +VH domains, opposite to their reversible folding ability. From the linear least square analysis of reversible folding ability and temperature-induced aggregation, we noticed that these two parameters are negatively correlated. VH domains having high reversible folding ability can refold after denaturation without becoming aggregated and thus highly soluble in solution. Alternatively, VH domains possessing less reversible folding ability become aggregated as they cannot refold after denaturation and thus less soluble in solution.

We measured the long-term solubility of the VH domains at room temperature. We observed a trend of room temperature solubility similar to their thermostability. From the linear least square analysis of room temperature solubility and thermostability, we observed a positive correlation between these two parameters. The high melting temperature (Tm) of a VH domain indicates its high stability and a stable VH domain is highly soluble at room temperature for a long time without getting aggregated. Conversely, VH domain with low melting temperature are less stable at room temperature and aggregated rapidly without denaturation and therefore less soluble in solution.

From the above analysis, we can conclude that the biophysical properties of VH domain such as thermostability, reversible folding ability, and solubility are not only determined by the net charge of the VH framework and corresponding CDR1 mutations. Rather, we observed position specific effects of lysine mutation in 31- and (31, 32)-positions of CDR1. Moreover, our designed +VH domain can better tolerate both negative and positive charge in CDR1, providing improved biophysical properties compared to VH-B1a.

Currently, we are working on the design and construction of phage-displayed VH domain libraries using VH-B1a and +VH domain as a template. We will combine charged mutations of VH framework and CDRs to get an improved VH scaffold for the construction of VH domain libraries. We will randomize both CDR1 and CDR3 of VH-B1a and +VH domain phage-displayed libraries and biopanning them against negatively charged DNA target (G-quadruplex and i-motif) that are associated with aging and cancer. CDR1 loop will be randomized with positively charged residues so that positively charged CDR1 loop can interact with negatively charged G-quadruplex and i-motif through electrostatic interaction, and help the CDR3 loop to bind with the target. Once we get any selective binder for G-quadruplex and i-motif from biopanning, then we will assess their binding affinity through Bio-layer interferometry assay. Finally, we will get VH domain against G-quadruplex and i-motif which are the therapeutic target in oncology.

7. REFERENCES

Akazawa-Ogawa, Y., Uegaki, K., and Hagihara, Y. (2016). The role of intra-domain disulfide bonds in heat-induced irreversible denaturation of camelid single domain VHH antibodies. J. Biochem. *159*, 111-121.

Alexander, A., Steinmetz, M., Barritault, D., Frangione, B., Franklin, E.C., Hood, L., and Buxbaum, J.N. (1982). Gamma Heavy chain disease in man: cDNA sequence supports partial gene deletion model. Proc. Natl. Acad. Sci. USA *79*, 3260-3264.

Andersen, D.C., and Reilly, D.E. (2004). Production technologies for monoclonal antibodies and their fragments. Curr. Opin. Biotechnol. *15*, 456-462.

Arbabi-Ghahroudi, M., To, R., Gaudette, N., Hirama, T., Ding, W., MacKenzie, R., and Tanha, J. (2009). Aggregation-resistant VHs selected by in vitro evolution tend to have disulfide-bonded loops and acidic isoelectric points. Protein Eng. Des. Sel. *22*, 59-66.

Baral, T.N., and Arbabi-Ghahroudi, M. (2012). Expression of single-domain antibodies in bacterial systems. Methods Mol. Biol. *911*, 257-275.

Baral, T.N., Magez, S., Stijlemans, B., Conrath, K., Vanhollebeke, B., Pays, E., Muyldermans, S., and De Baetselier, P. (2006). Experimental therapy of African trypanosomiasis with a nanobody-conjugated human trypanolytic factor. Nat. Med. *12*, 580-584.

Barthelemy, P.A., Raab, H., Appleton, B.A., Bond, C.J., Wu, P., Wiesmann, C., and Sidhu, S.S. (2008). Comprehensive analysis of the factors contributing to the stability and solubility of autonomous human VH domains. J. Biol. Chem. *283*, 3639-3654.

Bartunek, J., Barbato, E., Heyndrickx, G., Vanderheyden, M., Wijns, W., and Holz, J.B. (2013). Novel antiplatelet agents: ALX-0081, a Nanobody directed towards von Willebrand factor. J. Cardiovasc. Transl. Res. *6*, 355-363.

Betz, S.F. (1993). Disulfide bonds and the stability of globular proteins. Protein Sci. *2*, 1551-1558.

Bird, R.E., Hardman, K.D., Jacobson, J.W., Johnson, S., Kaufman, B.M., Lee, S.M., Lee, T., Pope, S.H., Riordan, G.S., and Whitlow, M. (1988). Single-chain antigen-binding proteins. Science *242*, 423-426.

Bond, C.J., Marsters, J.C., and Sidhu, S.S. (2003). Contributions of CDR3 to VHH domain stability and the design of monobody scaffolds for naive antibody libraries. J. Mol. Biol. *332*, 643-655.

Capriotti, E., Fariselli, P., Rossi, I., and Casadio, R. (2008). A three-state prediction of single point mutations on protein stability changes. BMC Bioinformatics *9 Suppl 2*, S6.

Chan, P.H., Pardon, E., Menzer, L., De Genst, E., Kumita, J.R., Christodoulou, J., Saerens, D., Brans, A., Bouillenne, F., Archer, D.B., Robinson, C.V., Muyldermans, S., Matagne, A., Redfield, C., Wyns, L., Dobson, C.M., and Dumoulin, M. (2008). Engineering a camelid

antibody fragment that binds to the active site of human lysozyme and inhibits its conversion into amyloid fibrils. Biochemistry *47*, 11041-11054.

Chothia, C., Lesk, A.M., Tramontano, A., Levitt, M., Smith-Gill, S.J., Air, G., Sheriff, S., Padlan, E.A., Davies, D., Tulip, W.R., and et al. (1989). Conformations of immunoglobulin hypervariable regions. Nature *342*, 877-883.

Ciaccio, N.A., and Laurence, J.S. (2009). Effects of disulfide bond formation and protein helicity on the aggregation of activating transcription factor 5. Mol. Pharm. *6*, 1205-1215.

Cortez-Retamozo, V., Backmann, N., Senter, P.D., Wernery, U., De Baetselier, P., Muyldermans, S., and Revets, H. (2004). Efficient cancer therapy with a nanobody-based conjugate. Cancer Res. *64*, 2853-2857.

Cortez-Retamozo, V., Lauwereys, M., Hassanzadeh Gh, G., Gobert, M., Conrath, K., Muyldermans, S., De Baetselier, P., and Revets, H. (2002). Efficient tumor targeting by single-domain antibody fragments of camels. Int. J. Cancer *98*, 456-462.

Das, R., and Baker, D. (2008). Macromolecular modeling with rosetta. Annu. Rev. Biochem. *77*, 363-382.

Davies, D.R., and Cohen, G.H. (1996). Interactions of protein antigens with antibodies. Proc. Natl. Acad. Sci. USA *93*, 7-12.

Davies, J., and Riechmann, L. (1994). 'Camelising' human antibody fragments: NMR studies on VH domains. FEBS Lett. *339*, 285-290.

Davies, J., and Riechmann, L. (1995). Antibody VH domains as small recognition units. Biotechnology (N. Y.) *13*, 475-479.

Davies, J., and Riechmann, L. (1996). Single antibody domains as small recognition units: design and in vitro antigen selection of camelized, human VH domains with improved protein stability. Protein Eng. *9*, 531-537.

de Marco, A. (2011). Biotechnological applications of recombinant single-domain antibody fragments. Microb. Cell Fact. *10*, 44.

Decanniere, K., Desmyter, A., Lauwereys, M., Ghahroudi, M.A., Muyldermans, S., and Wyns, L. (1999). A single-domain antibody fragment in complex with RNase A: non-canonical loop structures and nanomolar affinity using two CDR loops. Structure *7*, 361-370.

Deckers, N., Saerens, D., Kanobana, K., Conrath, K., Victor, B., Wernery, U., Vercruysse, J., Muyldermans, S., and Dorny, P. (2009). Nanobodies, a promising tool for species-specific diagnosis of *Taenia solium* cysticercosis. Int. J. Parasitol. *39*, 625-633.

Deffar, K., Shi, H., Li, L., Wang, X., and Zhu, X. (2009). Nanobodies-the new concept in antibody engineering. African journal of biotechnology *8*, 2645-2652.

Desmyter, A., Decanniere, K., Muyldermans, S., and Wyns, L. (2001). Antigen specificity and high affinity binding provided by one single loop of a camel single-domain antibody. J. Biol. Chem. *276*, 26285-26290.

Desmyter, A., Spinelli, S., Payan, F., Lauwereys, M., Wyns, L., Muyldermans, S., and Cambillau, C. (2002). Three camelid VHH domains in complex with porcine pancreatic alpha-amylase. Inhibition and versatility of binding topology. J. Biol. Chem. *277*, 23645-23650.

Desmyter, A., Transue, T.R., Ghahroudi, M.A., Thi, M.H., Poortmans, F., Hamers, R., Muyldermans, S., and Wyns, L. (1996). Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. Nat. Struct. Biol. *3*, 803-811.

Dimitrov, D.S. (2012). Therapeutic proteins. Methods in molecular biology (Clifton, N.J.) *899*, 26.

Dong, X., Stothard, P., Forsythe, I.J., and Wishart, D.S. (2004). PlasMapper: a web server for drawing and auto-annotating plasmid maps. Nucleic Acids Res. *32*, W660-664.

Dudgeon, K., Famm, K., and Christ, D. (2009). Sequence determinants of protein aggregation in human VH domains. Protein Eng. Des. Sel. *22*, 217-220.

Dudgeon, K., Rouet, R., Kokmeijer, I., Schofield, P., Stolp, J., Langley, D., Stock, D., and Christ, D. (2012). General strategy for the generation of human antibody variable domains with increased aggregation resistance. Proc. Natl. Acad. Sci. USA *109*, 10879-10884.

Elbing, K.L., and Brent, R. (2001). Media preparation and bacteriological tools. Curr. Protoc. Protein Sci. *Appendix 4*, Appendix 4A.

Els Conrath, K., Lauwereys, M., Wyns, L., and Muyldermans, S. (2001). Camel single-domain antibodies as modular building units in bispecific and bivalent antibody constructs. J. Biol. Chem. *276*, 7346-7350.

Ewert, S., Cambillau, C., Conrath, K., and Pluckthun, A. (2002). Biophysical properties of camelid VHH domains compared to those of human VH3 domains. Biochemistry *41*, 3628-3636.

Feldhaus, M.J., and Siegel, R.W. (2004). Yeast display of antibody fragments: a discovery and characterization platform. J. Immunol. Methods *290*, 69-80.

Fellouse, F.A., Barthelemy, P.A., Kelley, R.F., and Sidhu, S.S. (2006). Tyrosine plays a dominant functional role in the paratope of a synthetic antibody derived from a four amino acid code. J. Mol. Biol. *357*, 100-114.

Fellouse, F.A., Li, B., Compaan, D.M., Peden, A.A., Hymowitz, S.G., and Sidhu, S.S. (2005). Molecular recognition by a binary code. J. Mol. Biol. *348*, 1153-1162.

Fersht, A.R. (1997). Nucleation mechanisms in protein folding. Curr. Opin. Struct. Biol. *7*, 3-9.

Fraczkiewicz, R., and W. Braun. (1998a). A new efficient algorithm for calculating solvent accessible surface areas of macromolecules. Journal of Computational Chemistry *19*, 319-326.

Fraczkiewicz, R., and Werner Braun. (1998b). Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. Journal of computational chemistry *19*, 319-333.

Ghetie, V., Hubbard, J.G., Kim, J.K., Tsen, M.F., Lee, Y., and Ward, E.S. (1996). Abnormally short serum half-lives of IgG in beta 2-microglobulin-deficient mice. Eur. J. Immunol. *26*, 690-696.

Gong, R., Vu, B.K., Feng, Y., Prieto, D.A., Dyba, M.A., Walsh, J.D., Prabakaran, P., Veenstra, T.D., Tarasov, S.G., Ishima, R., and Dimitrov, D.S. (2009). Engineered human antibody constant domains with increased stability. J. Biol. Chem. *284*, 14203-14210.

Goto, Y., and Hamaguchi, K. (1979). The role of the intrachain disulfide bond in the conformation and stability of the constant fragment of the immunoglobulin light chain. J. Biochem. *86*, 1433-1441.

Govaert, J., Pellis, M., Deschacht, N., Vincke, C., Conrath, K., Muyldermans, S., and Saerens, D. (2012). Dual beneficial effect of interloop disulfide bond for single domain antibody fragments. J. Biol. Chem. *287*, 1970-1979.

Greenberg, A.S., Avila, D., Hughes, M., Hughes, A., McKinney, E.C., and Flajnik, M.F. (1995). A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. Nature *374*, 168-173.

Greenfield, N., and Fasman, G.D. (1969). Computed circular dichroism spectra for the evaluation of protein conformation. Biochemistry *8*, 4108-4116.

Greenfield, N.J. (2006a). Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. Nat. Protoc. *1*, 2527-2535.

Greenfield, N.J. (2006b). Using circular dichroism spectra to estimate protein secondary structure. Nat. Protoc. *1*, 2876-2890.

Greenwood, J., Clark, M., and Waldmann, H. (1993). Structural motifs involved in human IgG antibody effector functions. Eur. J. Immunol. *23*, 1098-1104.

Hagihara, Y., Mine, S., and Uegaki, K. (2007). Stabilization of an immunoglobulin fold domain by an engineered disulfide bond at the buried hydrophobic region. J. Biol. Chem. *282*, 36489-36495.

Hagihara, Y., and Saerens, D. (2014). Engineering disulfide bonds within an antibody. Biochim. Biophys. Acta. *1844*, 2016-2023.

Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Songa, E.B., Bendahman, N., and Hamers, R. (1993). Naturally occurring antibodies devoid of light chains. Nature *363*, 446-448.

Harmsen, M.M., Ruuls, R.C., Nijman, I.J., Niewold, T.A., Frenken, L.G., and de Geus, B. (2000). Llama heavy-chain V regions consist of at least four distinct subfamilies revealing novel sequence features. Mol. Immunol. *37*, 579-590.

Harmsen, M.M., van Solt, C.B., van Zijderveld-van Bemmel, A.M., Niewold, T.A., and van Zijderveld, F.G. (2006). Selection and optimization of proteolytically stable llama single-domain antibody fragments for oral immunotherapy. Appl. Microbiol. Biotechnol. *72*, 544-551.

Helms, L.R., and Wetzel, R. (1995). Destabilizing loop swaps in the CDRs of an immunoglobulin VL domain. Protein Sci. *4*, 2073-2081.

Hinds, K.R., and Litman, G.W. (1986). Major reorganization of immunoglobulin VH segmental elements during vertebrate evolution. Nature *320*, 546-549.

Hmila, I., Abdallah, R.B., Saerens, D., Benlasfar, Z., Conrath, K., Ayeb, M.E., Muyldermans, S., and Bouhaouala-Zahar, B. (2008). VHH, bivalent domains and chimeric Heavy chain-only antibodies with high neutralizing efficacy for scorpion toxin AahI'. Mol. Immunol. *45*, 3847-3856.

Holliger, P., Prospero, T., and Winter, G. (1993). "Diabodies": small bivalent and bispecific antibody fragments. Proc. Natl. Acad. Sci. USA *90*, 6444-6448.

Holt, L.J., Herring, C., Jespers, L.S., Woolven, B.P., and Tomlinson, I.M. (2003). Domain antibodies: proteins for therapy. Trends Biotechnol. *21*, 484-490.

Holzwarth, G., and Doty, P. (1965). The ultraviolet circular dichroism of polypeptides. J. Am. Chem. Soc. *87*, 218-228.

Howard, G.C., and Kaser, M.R. (2007). Making and using antibodies: a practical handbook. CRC press, Boca Raton, FL, 394.

Huang, L., Gainkam, L.O., Caveliers, V., Vanhove, C., Keyaerts, M., De Baetselier, P., Bossuyt, A., Revets, H., and Lahoutte, T. (2008). SPECT imaging with 99mTc-labeled EGFR-specific nanobody for *in vivo* monitoring of EGFR expression. Mol. Imaging Biol. *10*, 167-175.

Hussack, G., Hirama, T., Ding, W., Mackenzie, R., and Tanha, J. (2011). Engineered single-domain antibodies with high protease resistance and thermal stability. PLoS One *6*, e28218.

Hussack, G., Mackenzie, C.R., and Tanha, J. (2012). Characterization of single-domain antibodies with an engineered disulfide bond. Methods Mol. Biol. *911*, 417-429.

Huston, J.S., Levinson, D., Mudgett-Hunter, M., Tai, M.S., Novotný, J., Margolies, M.N., Ridge, R.J., Bruccoleri, R.E., Haber, E., and Crea, R. (1988). Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. Proc. Natl. Acad. Sci. USA *85*, 5879-5883.

Huston, J.S., Mudgett-Hunter, M., Tai, M.S., McCartney, J., Warren, F., Haber, E., and Oppermann, H. (1991). Protein engineering of single-chain Fv analogs and fusion proteins. Methods Enzymol. *203*, 46-88.

Hwang, W.Y., and Foote, J. (2005). Immunogenicity of engineered antibodies. Methods *36*, 3-10.

Ignatovich, O., Jespers, L., Tomlinson, I.M., and de Wildt, R.M. (2012). Creation of the large and highly functional synthetic repertoire of human VH and Vκ domain antibodies. Methods Mol. Biol. *911*, 39-63.

Inbar, D., Hochman, J., and Givol, D. (1972). Localization of antibody-combining sites within the variable portions of heavy and light chains. Proc. Natl. Acad. Sci. USA *69*, 2659-2662.

Ionescu, R.M., Vlasak, J., Price, C., and Kirchmeier, M. (2008). Contribution of variable domains to the stability of humanized IgG1 monoclonal antibodies. J. Pharm. Sci. *97*, 1414-1426.

Jespers, L., Schon, O., Famm, K., and Winter, G. (2004). Aggregation-resistant domain antibodies selected on phage by heat denaturation. Nat. Biotechnol. *22*, 1161-1165.

Jung, S., Honegger, A., and Pluckthun, A. (1999). Selection for improved protein stability by phage display. J. Mol. Biol. *294*, 163-180.

Kabat, E.A., Te Wu, T., Foeller, C., Perry, H.M., and Gottesman, K.S. (1992). Sequences of proteins of immunological interest. DIANE publishing.

Kim, D.Y., Kandalaft, H., Ding, W., Ryan, S., van Faassen, H., Hirama, T., Foote, S.J., MacKenzie, R., and Tanha, J. (2012). Disulfide linkage engineering for improving biophysical properties of human VH domains. Protein. Eng. Des. Sel. *25*, 581-589.

Koide, S. (2009). Engineering of recombinant crystallization chaperones. Curr. Opin. Struct. Biol. *19*, 449-457.

Kunkel, T.A., Roberts, J.D., and Zakour, R.A. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. *154*, 367-382.

Kvam, E., Sierks, M.R., Shoemaker, C.B., and Messer, A. (2010). Physico-chemical determinants of soluble intrabody expression in mammalian cell cytoplasm. Protein Eng. Des. Sel. *23*, 489-498.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680-685.

Lawrence, M.S., Phillips, K.J., and Liu, D.R. (2007). Supercharging proteins can impart unusual resilience. J. Am. Chem. Soc. *129*, 10110-10112.

Lee, C.C., Perchiacca, J.M., and Tessier, P.M. (2013). Toward aggregation-resistant antibodies by design. Trends Biotechnol. *31*, 612-620.

Lee, C.M., Iorno, N., Sierro, F., and Christ, D. (2007). Selection of human antibody fragments by phage display. Nat. Protoc. *2*, 3001-3008.

Lee, C.V., Liang, W.C., Dennis, M.S., Eigenbrot, C., Sidhu, S.S., and Fuh, G. (2004). High-affinity human antibodies from phage-displayed synthetic Fab libraries with a single framework scaffold. J. Mol. Biol. *340*, 1073-1093.

Liu, J., and Geyer, R.C. (2015). Engineering antibody Fragments for Intracellular applications. CRC press *2*, 521-544.

Ma, X., Barthelemy, P.A., Rouge, L., Wiesmann, C., and Sidhu, S.S. (2013). Design of synthetic autonomous VH domain libraries and structural analysis of a VH domain bound to vascular endothelial growth factor. J. Mol. Biol. *425*, 2247-2259.

Maass, D.R., Sepulveda, J., Pernthaner, A., and Shoemaker, C.B. (2007). Alpaca (Lama pacos) as a convenient source of recombinant camelid heavy chain antibodies (VHHs). J. Immunol. Methods *324*, 13-25.

Mandrup, O.A., Friis, N.A., Lykkemark, S., Just, J., and Kristensen, P. (2013). A novel heavy domain antibody library with functionally optimized complementarity determining regions. PLoS One *8*, e76834.

Mansfeld, J., Vriend, G., Dijkstra, B.W., Veltman, O.R., Van den Burg, B., Venema, G., Ulbrich-Hofmann, R., and Eijsink, V.G. (1997). Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond. J. Biol. Chem. *272*, 11152-11156.

Mason, J.M., Gibbs, N., Sessions, R.B., and Clarke, A.R. (2002). The influence of intramolecular bridges on the dynamics of a protein folding reaction. Biochemistry *41*, 12093-12099.

Matsumura, M., Signor, G., and Matthews, B.W. (1989). Substantial increase of protein stability by multiple disulphide bonds. Nature *342*, 291-293.

McCafferty, J., Griffiths, A.D., Winter, G., and Chiswell, D.J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. Nature *348*, 552-554.

Mian, I.S., Bradwell, A.R., and Olson, A.J. (1991). Structure, function and properties of antibody binding sites. J. Mol. Biol. *217*, 133-151.

Miklos, A.E., Kluwe, C., Der, B.S., Pai, S., Sircar, A., Hughes, R.A., Berrondo, M., Xu, J., Codrea, V., Buckley, P.E., Calm, A.M., Welsh, H.S., Warner, C.R., Zacharko, M.A., Carney, J.P., Gray, J.J., Georgiou, G., Kuhlman, B., and Ellington, A.D. (2012). Structure-based design of supercharged, highly thermoresistant antibodies. Chem. Biol. *19*, 449-455.

Morrison, F.A. (2014). Obtaining uncertainty measures on slope and intercept of a least squares fit with Excel's LINEST. Department of Chemical Engineering, Michigan Technological University, Houghton, Michigan, USA *25*.

Muyldermans, S. (2001). Single domain camel antibodies: current status. J. Biotechnol. *74*, 277-302.

Muyldermans, S. (2013). Nanobodies: natural single-domain antibodies. Annu. Rev. Biochem. *82*, 775-797.

Muyldermans, S., Atarhouch, T., Saldanha, J., Barbosa, J.A., and Hamers, R. (1994). Sequence and structure of VH domain from naturally occurring camel heavy chain immunoglobulins lacking light chains. Protein Eng. *7*, 1129-1135.

Muyldermans, S., Cambillau, C., and Wyns, L. (2001). Recognition of antigens by single-domain antibody fragments: the superfluous luxury of paired domains. Trends Biochem. Sci. *26*, 230-235.

Nguyen, V.K., Hamers, R., Wyns, L., and Muyldermans, S. (2000). Camel heavy-chain antibodies: diverse germline VHH and specific mechanisms enlarge the antigen-binding repertoire. EMBO J. *19*, 921-930.

Nuttall, S.D. (2012). Overview and discovery of IgNARs and generation of VNARs. Methods Mol. Biol. *911*, 27-36.

Nuttall, S.D., Humberstone, K.S., Krishnan, U.V., Carmichael, J.A., Doughty, L., Hattarki, M., Coley, A.M., Casey, J.L., Anders, R.F., Foley, M., Irving, R.A., and Hudson, P.J. (2004). Selection and affinity maturation of IgNAR variable domains targeting *Plasmodium falciparum* AMA1. Proteins *55*, 187-197.

Nuttall, S.D., Krishnan, U.V., Doughty, L., Pearson, K., Ryan, M.T., Hoogenraad, N.J., Hattarki, M., Carmichael, J.A., Irving, R.A., and Hudson, P.J. (2003). Isolation and characterization of an IgNAR variable domain specific for the human mitochondrial translocase receptor Tom70. Eur. J. Biochem. *270*, 3543-3554.

Pace, C.N., and J. Martin Scholtz. (1997). Measuring the conformational stability of a protein. Protein structure: A practical approach *2*, 299-321.

Padlan, E.A. (1994). Anatomy of the antibody molecule. Mol. Immunol. *31*, 169-217.

Pal, G., Kouadio, J.L., Artis, D.R., Kossiakoff, A.A., and Sidhu, S.S. (2006). Comprehensive and quantitative mapping of energy landscapes for protein-protein interactions by rapid combinatorial scanning. J. Biol. Chem. *281*, 22378-22385.

Perchiacca, J.M., Bhattacharya, M., and Tessier, P.M. (2011). Mutational analysis of domain antibodies reveals aggregation hotspots within and near the complementarity determining regions. Proteins *79*, 2637-2647.

Perchiacca, J.M., Ladiwala, A.R., Bhattacharya, M., and Tessier, P.M. (2012). Aggregation-resistant domain antibodies engineered with charged mutations near the edges of the complementarity-determining regions. Protein Eng. Des. Sel. *25*, 591-601.

Perchiacca, J.M., Lee, C.C., and Tessier, P.M. (2014). Optimal charged mutations in the complementarity-determining regions that prevent domain antibody aggregation are dependent on the antibody scaffold. Protein Eng. Des. Sel. *27*, 29-39.

Perchiacca, J.M., and Tessier, P.M. (2012). Engineering aggregation-resistant antibodies. Annu. Rev. Chem. Biomol. Eng. *3*, 263-286.

Perez, J.M., Renisio, J.G., Prompers, J.J., van Platerink, C.J., Cambillau, C., Darbon, H., and Frenken, L.G. (2001). Thermal unfolding of a llama antibody fragment: a two-state reversible process. Biochemistry *40*, 74-83.

Plückthun, A., and Pack, P. (1997). New protein engineering approaches to multivalent and bispecific antibody fragments. Immunotechnology *3*, 83-105.

Porter, R.R. (1958). Separation and isolation of fractions of rabbit gamma-globulin containing the antibody and antigenic combining sites. Nature *182*, 670-671.

Porter, R.R. (1959). The hydrolysis of rabbit γ-globulin and antibodies with crystalline papain. Biochem. J. *73*, 119-126.

Proba, K., Worn, A., Honegger, A., and Pluckthun, A. (1998). Antibody scFv fragments without disulfide bonds made by molecular evolution. J. Mol. Biol. *275*, 245-253.

Ries, J., Kaplan, C., Platonova, E., Eghlidi, H., and Ewers, H. (2012). A simple, versatile method for GFP-based super-resolution microscopy via nanobodies. Nat. Methods *9*, 582-584.

Rinaldi, A.S., Freund, G., Desplancq, D., Sibler, A.P., Baltzinger, M., Rochel, N., Mely, Y., Didier, P., and Weiss, E. (2013). The use of fluorescent intrabodies to detect endogenous gankyrin in living cancer cells. Exp. Cell Res. *319*, 838-849.

Roovers, R.C., Laeremans, T., Huang, L., De Taeye, S., Verkleij, A.J., Revets, H., de Haard, H.J., and van Bergen en Henegouwen, P.M. (2007). Efficient inhibition of EGFR signaling and of tumour growth by antagonistic anti-EFGR Nanobodies. Cancer Immunol. Immunother. *56*, 303-317.

Rosebrough, S.F. (1993). Pharmacokinetics and biodistribution of radiolabeled avidin, streptavidin and biotin. Nucl. Med. Biol. *20*, 663-668.

Rothbauer, U., Zolghadr, K., Muyldermans, S., Schepers, A., Cardoso, M.C., and Leonhardt, H. (2008). A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. Mol. Cell Proteomics *7*, 282-289.

Rothbauer, U., Zolghadr, K., Tillib, S., Nowak, D., Schermelleh, L., Gahl, A., Backmann, N., Conrath, K., Muyldermans, S., Cardoso, M.C., and Leonhardt, H. (2006). Targeting and tracing antigens in live cells with fluorescent nanobodies. Nat. Methods *3*, 887-889.

Rothlisberger, D., Honegger, A., and Pluckthun, A. (2005). Domain interactions in the Fab fragment: a comparative evaluation of the single-chain Fv and Fab format engineered with variable domains of different stability. J. Mol. Biol. *347*, 773-789.

Rouet, R., Dudgeon, K., and Christ, D. (2012). Generation of human single domain antibody repertoires by Kunkel mutagenesis. Methods Mol. Biol. *907*, 195-209.

Roux, K.H., Greenberg, A.S., Greene, L., Strelets, L., Avila, D., McKinney, E.C., and Flajnik, M.F. (1998). Structural analysis of the nurse shark (new) antigen receptor (NAR): molecular convergence of NAR and unusual mammalian immunoglobulins. Proc. Natl. Acad. Sci. USA *95*, 11804-11809.

Saerens, D., Conrath, K., Govaert, J., and Muyldermans, S. (2008a). Disulfide bond introduction for general stabilization of immunoglobulin heavy-chain variable domains. J. Mol. Biol. *377*, 478-488.

Saerens, D., Frederix, F., Reekmans, G., Conrath, K., Jans, K., Brys, L., Huang, L., Bosmans, E., Maes, G., Borghs, G., and Muyldermans, S. (2005). Engineering camel single-domain antibodies and immobilization chemistry for human prostate-specific antigen sensing. Anal. Chem. *77*, 7547-7555.

Saerens, D., Ghassabeh, G.H., and Muyldermans, S. (2008b). Single-domain antibodies as building blocks for novel therapeutics. Curr. Opin. Pharmacol. *8*, 600-608.

Saerens, D., Kinne, J., Bosmans, E., Wernery, U., Muyldermans, S., and Conrath, K. (2004). Single domain antibodies derived from dromedary lymph node and peripheral blood lymphocytes sensing conformational variants of prostate-specific antigen. J. Biol. Chem. *279*, 51965-51972.

Shire, S.J. (2009). Formulation and manufacturability of biologics. Curr. Opin. Biotechnol. *20*, 708-714.

Shire, S.J., Shahrokh, Z., and Liu, J. (2004). Challenges in the development of high protein concentration formulations. J. Pharm. Sci. *93*, 1390-1402.

Sidhu, S.S., Li, B., Chen, Y., Fellouse, F.A., Eigenbrot, C., and Fuh, G. (2004). Phage-displayed antibody libraries of synthetic heavy chain complementarity determining regions. J. Mol. Biol. *338*, 299-310.

Sidhu, S.S., Weiss, G.A., and Wells, J.A. (2000). High copy display of large proteins on phage for functional selections. J. Mol. Biol. *296*, 487-495.

Smith, G.P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science *228*, 1315-1317.

Spinelli, S., Frenken, L., Bourgeois, D., de Ron, L., Bos, W., Verrips, T., Anguille, C., Cambillau, C., and Tegoni, M. (1996). The crystal structure of a llama heavy chain variable domain. Nat. Struct. Biol. *3*, 752-757.

Spinelli, S., Frenken, L.G., Hermans, P., Verrips, T., Brown, K., Tegoni, M., and Cambillau, C. (2000). Camelid heavy-chain variable domains provide efficient combining sites to haptens. Biochemistry *39*, 1217-1222.

Spinelli, S., Tegoni, M., Frenken, L., van Vliet, C., and Cambillau, C. (2001). Lateral recognition of a dye hapten by a llama VHH domain. J. Mol. Biol. *311*, 123-129.

Stijlemans, B., Conrath, K., Cortez-Retamozo, V., Van Xong, H., Wyns, L., Senter, P., Revets, H., De Baetselier, P., Muyldermans, S., and Magez, S. (2004). Efficient targeting of conserved cryptic epitopes of infectious agents by single domain antibodies. African trypanosomes as paradigm. J. Biol. Chem. *279*, 1256-1261.

Streltsov, V.A., Carmichael, J.A., and Nuttall, S.D. (2005). Structure of a shark IgNAR antibody variable domain and modeling of an early-developmental isotype. Protein Sci. *14*, 2901-2909.

Sundaresan, G., Yazaki, P.J., Shively, J.E., Finn, R.D., Larson, S.M., Raubitschek, A.A., Williams, L.E., Chatziioannou, A.F., Gambhir, S.S., and Wu, A.M. (2003). 124I-labeled engineered anti-CEA minibodies and diabodies allow high-contrast, antigen-specific small-animal PET imaging of xenografts in athymic mice. J. Nucl. Med. *44*, 1962-1969.

Sundberg, E.J., Urrutia, M., Braden, B.C., Isern, J., Tsuchiya, D., Fields, B.A., Malchiodi, E.L., Tormo, J., Schwarz, F.P., and Mariuzza, R.A. (2000). Estimation of the hydrophobic effect in an antigen-antibody protein-protein interface. Biochemistry *39*, 15375-15387.

Tan, P., Mitchell, D.A., Buss, T.N., Holmes, M.A., Anasetti, C., and Foote, J. (2002). "Superhumanized" antibodies: reduction of immunogenic potential by complementarity-determining region grafting with human germline sequences: application to an anti-CD28. J. Immunol. *169*, 1119-1125.

Tonikian, R., and Sidhu, S.S. (2012). Selecting and purifying autonomous human variable heavy (VH) domains. Methods Mol. Biol. *911*, 327-353.

Van de Broek, B., Devoogdt, N., D'Hollander, A., Gijs, H.L., Jans, K., Lagae, L., Muyldermans, S., Maes, G., and Borghs, G. (2011). Specific cell targeting with nanobody conjugated branched gold nanoparticles for photothermal therapy. ACS Nano *5*, 4319-4328.

Venyaminov, S., Baikalov, I.A., Shen, Z.M., Wu, C.S., and Yang, J.T. (1993). Circular dichroic analysis of denatured proteins: inclusion of denatured proteins in the reference set. Anal. Biochem. *214*, 17-24.

Verma, R., Boleti, E., and George, A.J. (1998). Antibody engineering: comparison of bacterial, yeast, insect and mammalian expression systems. J. Immunol. Methods *216*, 165-181.

Vincke, C., Loris, R., Saerens, D., Martinez-Rodriguez, S., Muyldermans, S., and Conrath, K. (2009). General strategy to humanize a camelid single-domain antibody and identification of a universal humanized nanobody scaffold. J. Biol. Chem. *284*, 3273-3284.

Vincke, C., and Muyldermans, S. (2012). Introduction to heavy chain antibodies and derived nanobodies. Methods Mol. Biol. *911*, 15-26.

Walper, S.A., Liu, J.L., Zabetakis, D., Anderson, G.P., and Goldman, E.R. (2014). Development and evaluation of single domain antibodies for vaccinia and the L1 antigen. PLoS One *9*, e106263.

Ward, E.S., Güssow, D., Griffiths, A.D., Jones, P.T., and Winter, G. (1989). Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli*. Nature *341*, 544-546.

Weiss, G.A., Watanabe, C.K., Zhong, A., Goddard, A., and Sidhu, S.S. (2000). Rapid mapping of protein functional epitopes by combinatorial alanine scanning. Proc. Natl. Acad. Sci. USA *97*, 8950-8954.

Wetzel, R., Perry, L.J., Baase, W.A., and Becktel, W.J. (1988). Disulfide bonds and thermal stability in T4 lysozyme. Proc. Natl. Acad. Sci. USA *85*, 401-405.

Wozniak-Knopp, G., Stadlmann, J., and Ruker, F. (2012). Stabilisation of the Fc fragment of human IgG1 by engineered intradomain disulfide bonds. PLoS One *7*, e30083.

Young, N.M., MacKenzie, C.R., Narang, S.A., Oomen, R.P., and Baenziger, J.E. (1995). Thermal stabilization of a single-chain Fv antibody fragment by introduction of a disulphide bond. FEBS Lett. *377*, 135-139.

Zalevsky, J., Chamberlain, A.K., Horton, H.M., Karki, S., Leung, I.W., Sproule, T.J., Lazar, G.A., Roopenian, D.C., and Desjarlais, J.R. (2010). Enhanced antibody half-life improves *in vivo* activity. Nat. Biotechnol. *28*, 157-159.

Zuckier, L.S., and DeNardo, G.L. (1997). Trials and tribulations: oncological antibody imaging comes to the fore. Semin. Nucl. Med. *27*, 10-29.